



*“Impulsando la Sociedad del
Conocimiento”*



Manual de Usuario

ERLA

2011



DESARROLLADORES

Juan Carlos Buenaño Cordero

Celia De La Cruz Cedeño

Roger Fernández Hidalgo

Andrea Fuentes Pugilla

Sindy Macías Cabrera

Estefany Minalla Álava

Edgar Murillo Alvarado

César Pincay Chiquito

Raúl Pinos Loayza

Nathaly Rivera Flores

Mario Solórzano Carvajal

PROFESOR DE LA MATERIA DE GRADUACIÓN

Gaudencio Zurita Herrera

Contenido

1	INTRODUCCIÓN.....	1
	Objetivos	2
	Revisión general	2
	Convenciones generales de este manual	2
	Datos técnicos	3
	Caso de estudio	3
	Instalación	4
2	ENTORNO GRÁFICO	6
	Objetivos	6
	Revisión general	6
	Inicio de ERLA.....	7
	Barra de título.....	8
	Barra de menues	9
	Menú Archivo	9
	Menú Ver	10
	Menú Ventanas	13
	Menú Ayuda.....	15
	Barra de herramientas.....	15
	Barra de herramientas de la ventana principal	16
	Barra de herramientas de la ventana de datos.....	16
	Barra de herramientas de la ventana de reporte	17
	Ventana de datos.....	17
	Editor de variables	18
	Importar datos.....	21

Generar variables aleatorias.....	23
Exportar ventana de datos	25
Buscar y reemplazar	26
Ventana de reporte	27
Guardar una ventana de reporte.....	27
Abrir una ventana de reporte.....	28
Imprimir reporte.....	29
Ventana de gráficos	32
Barra de estado.....	34
Cuadro de diálogo.....	35
3 ESTADÍSTICA DESCRIPTIVA.....	37
Objetivos	37
Revisión general	37
Gráficos	38
Diagrama de barras	38
Histograma	39
Diagrama de cajas.....	40
Gráfico de dispersión.....	41
Matriz de dispersión.....	42
Tabla de frecuencia	43
Variable categórica.....	43
Variable cuantitativa	44
Tablas cruzadas	45
Estadísticas descriptivas	46
4 CORRELACIÓN, REGRESIÓN Y ANÁLISIS DE VARIANZA	48
Objetivos	48

Revisión general	48
Correlación.....	49
Regresión.....	51
Regresión lineal simple	51
Regresión lineal simple polinómica	51
Gráfico de regresión	55
Regresión lineal múltiple	56
Regresión lineal múltiple polinómica con interacciones	56
Gráfico de regresión	60
Estimaciones.....	61
Verificación de la calidad en modelos de regresión lineal.....	62
Verificación de supuestos del error	64
Puntos de Influencia y Valores Aberrantes	67
Multicolinealidad.....	69
Selección de modelos	70
Regresión logística	73
Regresión Poisson.....	76
Regresión Ridge	79
Regresión robusta.....	82
Análisis de varianza	84
Análisis de varianza de un factor	84
Análisis de varianza de dos o más factores	88
5 RECOMENDACIONES.....	92
6 REFERENCIAS BIBLIOGRÁFICAS	93
ANEXO I	94

Índice de figuras

Figura 1.1	<i>“Instalación de ERLA”</i>	5
Figura 2.1	<i>“Inicio ERLA”</i>	7
Figura 2.2	<i>“Ventana principal de ERLA”</i>	8
Figura 2.3	<i>“Barra de título de la ventana principal”</i>	8
Figura 2.4	<i>“Barra de Menues de la Ventana Principal”</i>	9
Figura 2.5	<i>“Menú Archivo”</i>	9
Figura 2.6	<i>“Menú Ver”</i>	10
Figura 2.7	<i>“Cuadro de diálogo Opciones del menú Ver”</i>	10
Figura 2.8	<i>“Cuadro de diálogo Color”</i>	11
Figura 2.9	<i>“Cuadro de diálogo Fuente”</i>	12
Figura 2.10	<i>“Menú Ventanas”</i>	13
Figura 2.11	<i>“Organizar Ventanas”</i>	14
Figura 2.12	<i>“Menú Ayuda”</i>	15
Figura 2.13	<i>“Barra de herramientas de la ventana principal”</i>	15
Figura 2.14	<i>“Ventana de datos”</i>	17
Figura 2.15	<i>“Editor de variables”</i>	18
Figura 2.16	<i>“Etiquetas de variables”</i>	19
Figura 2.17	<i>“Agregar una variable”</i>	19
Figura 2.18	<i>“Eliminar variables”</i>	20
Figura 2.19	<i>“Etiquetar una variable categórica”</i>	21
Figura 2.20	<i>“Importar datos desde un archivo *.csv”</i>	22
Figura 2.21	<i>“Funciones para generar números aleatorios”</i>	23
Figura 2.22	<i>“Exportar ventana de datos”</i>	25

Figura 2.23	<i>“Buscar o reemplazar”</i>	26
Figura 2.24	<i>“Ventana de reporte de ERLA”</i>	27
Figura 2.25	<i>“Guardar una ventana de reporte en ERLA”</i>	28
Figura 2.26	<i>“Abrir una ventana de reporte en ERLA”</i>	28
Figura 2.27	<i>“Vista previa de impresión en ERLA”</i>	29
Figura 2.28	<i>“Configurar impresión en ERLA”</i>	30
Figura 2.29	<i>“Ventana de gráficos”</i>	32
Figura 2.30	<i>“Imprimir un gráfico”</i>	33
Figura 2.31	<i>“Modificar un gráfico”</i>	34
Figura 2.32	<i>“Barra de estado”</i>	34
Figura 2.33	<i>“Cuadro de diálogo”</i>	35
Figura 3.1	<i>“Diagrama de Barras”</i>	38
Figura 3.2	<i>“Histograma”</i>	39
Figura 3.3	<i>“Diagrama de cajas”</i>	40
Figura 3.4	<i>“Gráfico de dispersión”</i>	41
Figura 3.5	<i>“Matriz de dispersión”</i>	42
Figura 3.6	<i>“Tabla de frecuencias de una variable categórica”</i>	43
Figura 3.7	<i>“Tabla de frecuencias de una variable cuantitativa”</i>	44
Figura 3.8	<i>“Tablas cruzadas”</i>	45
Figura 3.9	<i>“Estadísticas descriptivas”</i>	47
Figura 4.1	<i>“Matriz de correlación”</i>	50
Figura 4.2	<i>“Regresión lineal simple”</i>	52
Figura 4.3	<i>“Regresión lineal simple polinómica”</i>	54
Figura 4.4	<i>“Gráfico de regresión 2D”</i>	55
Figura 4.5	<i>“Regresión lineal múltiple”</i>	57
Figura 4.6	<i>“Generar una variable de interacción”</i>	58

Figura 4.7	<i>“Regresión lineal múltiple polinómica con interacciones”</i>	59
Figura 4.8	<i>“Gráfico de regresión 3D”</i>	60
Figura 4.9	<i>“Estimaciones en un modelo de regresión lineal”</i>	62
Figura 4.10	<i>“Medidas de calidad del modelo”</i>	63
Figura 4.11	<i>“Gráfico de normalidad del error”</i>	64
Figura 4.12	<i>“Gráfico de homocedasticidad”</i>	65
Figura 4.13	<i>“Gráfico de independencia de errores”</i>	66
Figura 4.14	<i>“Prueba Durbin-Watson”</i>	66
Figura 4.15	<i>“Apalancamiento”</i>	67
Figura 4.16	<i>“Distancia de Cook”</i>	67
Figura 4.17	<i>“Valores atípicos potenciales”</i>	68
Figura 4.18	<i>“Factor de agrandamiento de la varianza”</i>	69
Figura 4.19	<i>“Número de condición”</i>	69
Figura 4.20	<i>“Selección de Modelos”</i>	71
Figura 4.21	<i>“Regresión logística”</i>	73
Figura 4.22	<i>“Regresión logística: Resultados”</i>	74
Figura 4.23	<i>“Regresión Poisson”</i>	76
Figura 4.24	<i>“Regresión Poisson: Resultados”</i>	77
Figura 4.25	<i>“Regresión Ridge”</i>	80
Figura 4.26	<i>“Regresión robusta”</i>	82
Figura 4.27	<i>“Regresión robusta: Resultados”</i>	83
Figura 4.28	<i>“Ingreso de datos para el análisis de varianza de un factor”</i>	85
Figura 4.29	<i>“Análisis de varianza de un factor”</i>	86
Figura 4.30	<i>“Ingreso de datos para el análisis de varianza de dos o más factores”</i>	89
Figura 4.31	<i>“Análisis de varianza de dos o más factores”</i>	90

Índice de tablas

Tabla 1.1: “Ficha técnica del estudio: Imagen de la ESPOLE en Guayaquil”.	4
Tabla 2.1: “Lista de funciones predefinidas de la Calculadora”	24
Tabla 4.1: “Datos para regresión Ridge”	79
Tabla 4.2: “Datos para el análisis de varianza de una vía: Producción lechera”	84
Tabla 4.3: “Datos para el análisis de varianza de dos factores de la vida (en horas) del diseño de la batería”	88



INTRODUCCIÓN

ERLA es un software estadístico gratuito, especializado en la técnica de Regresión Lineal desarrollado por estudiantes del Instituto de Ciencias Matemáticas de la Escuela Superior Politécnica del Litoral mediante el uso del MCR¹ (MATLAB Component Runtime) y una interfaz gráfica desarrollada en .NET.

Este Software trabaja con funciones propias de MATLAB y otras funciones personalizadas para propósitos estadísticos y de ingeniería.

La información contenida en este manual muestra cómo generar gráficos estadísticos, obtener estadísticas descriptivas, tablas de frecuencias y modelos de regresión tales como: Regresión Lineal, Poisson, Logística, Ridge y Robusta. Además aprenderá a realizar un Análisis de Varianza de n vías.

Al ser ERLA, un software especializado en la técnica de Regresión Lineal, es posible evaluar la calidad de los modelos obtenidos, realizar estimaciones de todos los modelos que se hayan generado y además seleccionar el mejor modelo considerando todas las variables que usted considere sean relevantes en el estudio.

Para más información del Software y sus versiones visite: <http://code.google.com/p/erla>

¹ El MCR (*MATLAB Compiler Runtime*), es un conjunto independiente de librerías compartidas que permite la ejecución de archivos de MATLAB en computadores sin una versión instalada de MATLAB.

Objetivos

A través de este manual, usted aprenderá a:

- Instalar e iniciar la aplicación ERLA.
- Aplicar métodos y técnicas estadísticas utilizando ERLA.

Revisión general

Este manual consta de 6 Capítulos:

1. Introducción
2. Entorno Gráfico
3. Estadística Descriptiva
4. Correlación, Regresión y Análisis de Varianza
5. Recomendaciones
6. Referencias Bibliográficas

En cada uno de los capítulos 2, 3 y 4 se indican los objetivos del mismo junto a una revisión general de lo que se va a aprender.

Convenciones generales de este manual

[Enter]	Indica una tecla. Por ejemplo la tecla [Enter].
[Alt]+[D]	Indica que debe mantener presionada una tecla y luego presionar otra. Por ejemplo, mantenga presionada la tecla [Alt] y luego presione la tecla [D].
Archivo > Salir	Indica un comando del menú; en este caso, elija Salir en el menú Archivo .
Pulsar Aceptar .	El texto en negrita resalta nombres de ventanas, cuadros de diálogos, barras de herramientas botones y menues.
Ingrese <i>Texto 1</i> .	El texto en cursiva especifica el texto que debe ingresar en un cuadro de diálogo.



El icono que aparece a la izquierda le permitirá obtener información adicional sobre el tema desarrollado.

Datos técnicos

Nombre: ERLA (Estadística de Regresión Lineal Avanzada)

Plataforma de desarrollo: Microsoft Visual Basic .NET 2008, MATLAB 2010

Sistema operativo: Windows XP, Vista, 7

Idioma: Español

Año de desarrollo: 2010

Requisitos de instalación: Sistema operativo Windows con Microsoft .NET Framework 3.0 o superior, MATLAB Compiler Runtime 7.13 o superior.

Requisitos mínimos de sistema: Procesador Intel o AMD x86 con soporte SSE2, 300 MB disponibles en disco duro, 1024 MB de memoria RAM.

1.1 Caso de estudio

Para explicar cómo se realiza el análisis de datos en ERLA, se ha considerado, guardando la correspondiente confidencialidad que la ética estadística exige, una base de datos correspondiente a un estudio realizado en la Escuela Superior Politécnica del Litoral por el Centro de Estudios e Investigaciones Estadísticas, llamado “Imagen de la ESPOL en Guayaquil”. En la Tabla 1.1 se presenta la información Técnica del estudio. El formulario administrado a las unidades de investigación incluidas en la muestra se presenta en el ANEXO 1.

Este estudio se realizó con la finalidad de “medir” la Imagen de la ESPOL en la ciudad de Guayaquil y su posición en comparación con las demás universidades de la ciudad. Para obtener más información sobre los resultados de este estudio visite: www.icm.espol.edu.ec/estadisticas.

En este manual se han considerado un total de 388 registros, los cuales corresponden a las personas que conocían la ESPOL. Los 52 restantes la desconocían.

Tabla 1.1
ERLA 2011 Estadística ICM/ESPOL
"Ficha técnica del estudio: Imagen de la ESPOL en Guayaquil"

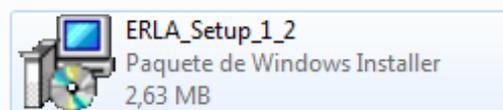
Población Objetivo:	Hogares de la ciudad Guayaquil
Tamaño de la Población Objetivo:	N = 393141
Fuente de Datos:	Encuesta, Tamaño de la muestra n = 440, Error de diseño E = 0.0467, Nivel de significancia $\alpha = 0.05$
Tipo de Muestreo:	Muestreo Aleatorio Bietápico, en el que la unidad primaria de investigación es el jefe del hogar o la persona mayor de edad que se encuentre en el hogar de la entrevista.
Fecha de realización del trabajo de campo:	Entre el 21 y 23 de enero de 2011

Instalación

Usted podrá obtener el instalador de ERLA a través del siguiente enlace:

<http://code.google.com/p/erla/>

Una vez obtenido el instalador de ERLA, pulse dos veces el archivo **ERLA_Setup_[versión].msi**



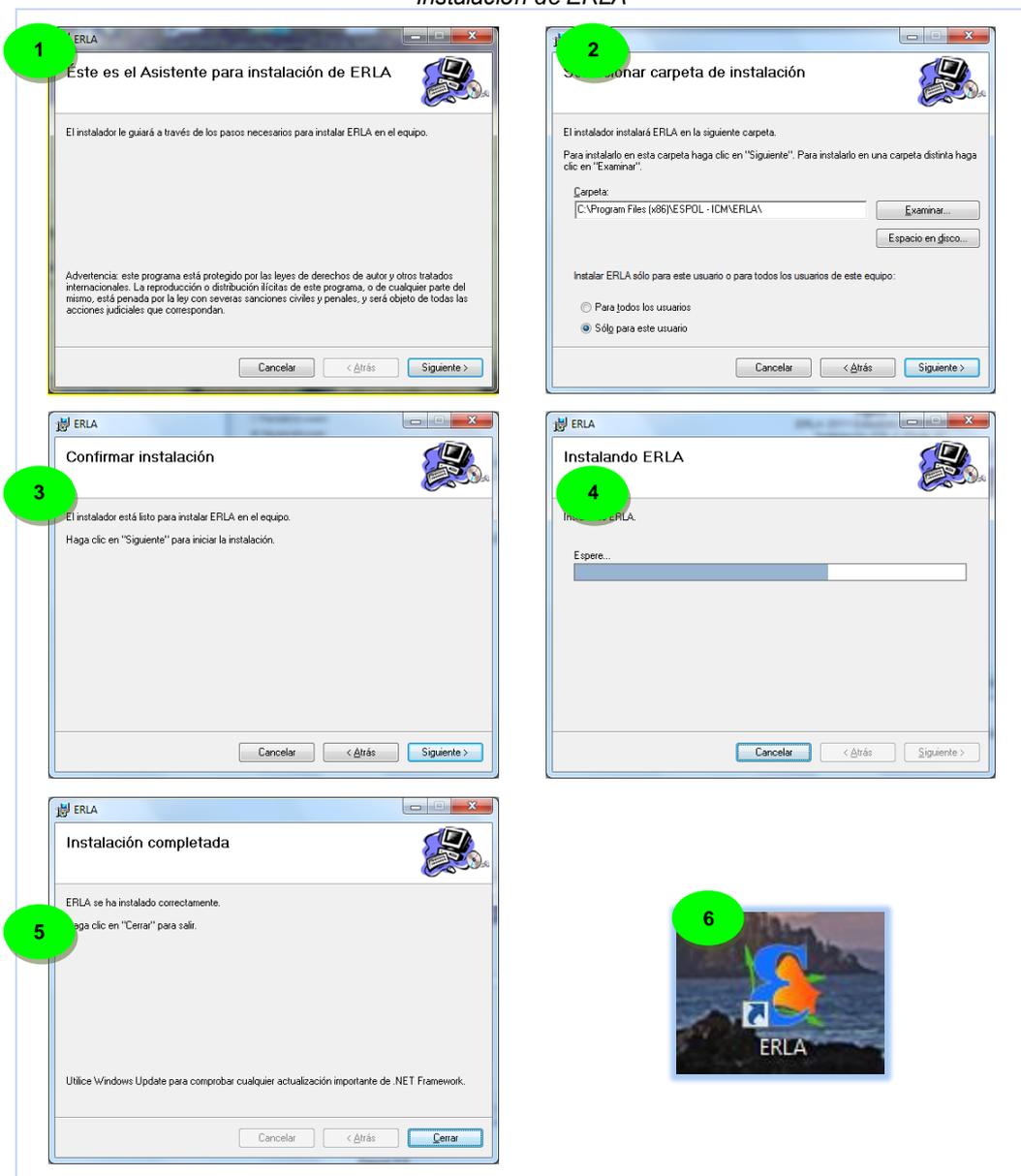
1. Lea el cuadro de dialogo del asistente de instalación y luego pulse el botón **Siguiente**.
2. Por defecto se instalará en la carpeta **Archivos de programa** de la unidad de sistema (C:).
3. Para iniciar la instalación pulse el botón **Siguiente**.
4. Espere mientras se realiza la instalación.
5. Una vez que se haya completado la instalación pulse el

botón **Cerrar**.

- Una vez instalado ERLA, se crea un acceso directo en el escritorio y otro en el menú **Inicio** a través de los cuales puede empezar a utilizar el software.

Véase Figura 1.1

Figura 1.1
ERLA 2011 Estadística ICM/ESPOL
"Instalación de ERLA"





ENTORNO GRÁFICO

Objetivos

- Familiarizarse con el ambiente de trabajo de ERLA.
- Ingresar, importar y exportar datos en ERLA.
- Abrir, guardar e imprimir reportes en ERLA.

Revisión general

En este capítulo se prepara al usuario para empezar a utilizar ERLA. Se revisan detenidamente cada uno de los menús y submenús, el tipo de variables que pueden definirse, cómo importar y exportar datos; y cómo guardar y abrir reportes. También se explica cómo deben seleccionarse las variables según la técnica estadística que vaya a usar en ERLA.

Inicio de ERLA

Luego de instalar ERLA, abra la **ventana principal** pulsando dos veces el acceso directo del escritorio o:

En la barra de tareas de Windows, elija **Inicio** > **Todos los Programas** > **ERLA** > **Software Estadístico ERLA**.

Mientras se carga el programa usted podrá visualizar la Figura 2.1.

Figura 2.1
ERLA 2011 Estadística ICM/ESPOL
"Inicio ERLA"

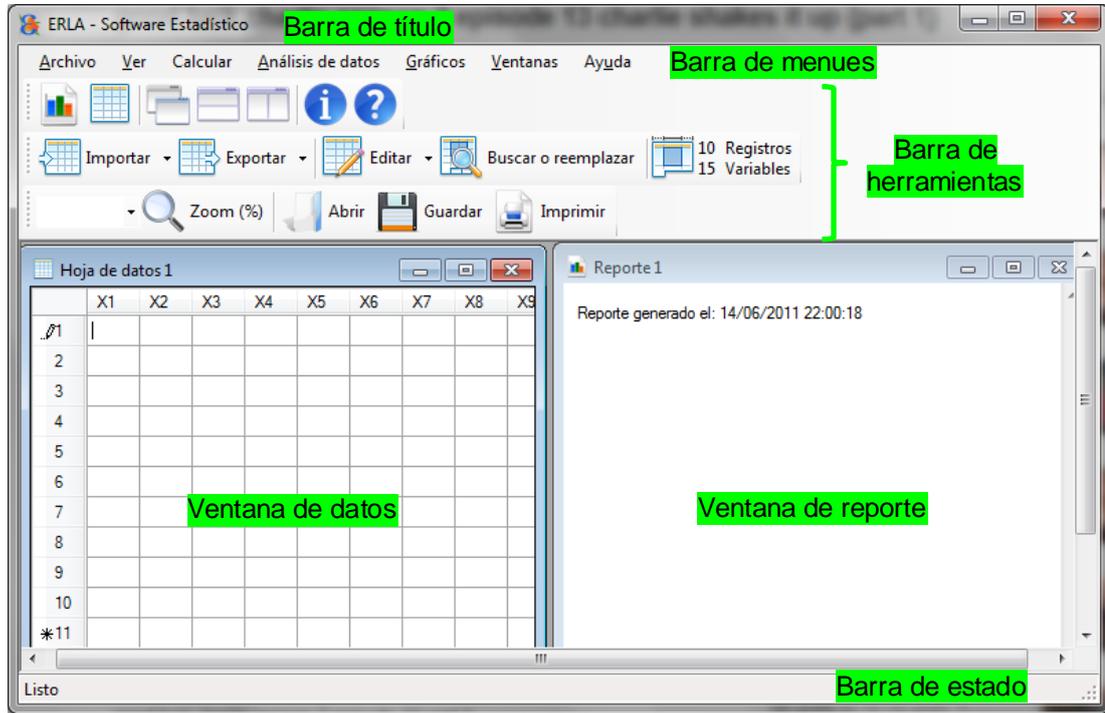


La **ventana principal** de ERLA está formada por una **barra de título**, una **barra de menues**, una **barra de herramientas** y además contiene dos ventanas:

- **Ventana de datos**, aquí se deben ingresar los datos correspondientes.
- **Ventana de reporte**, en la cual se presentan los resultados.

Por último en la parte inferior se encuentra la **barra de estado**. Véase Figura 2.2.

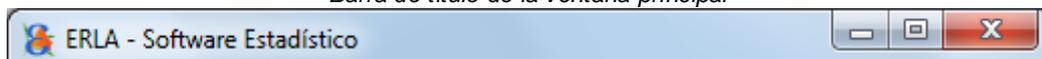
Figura 2.2
ERLA 2011 Estadística ICM/ESPOL
"Ventana principal de ERLA"



Barra de título

La barra de título está ubicada en la parte superior izquierda de la **ventana principal** y contiene el ícono y el nombre del Software en la parte derecha; además los botones **Minimizar** , **Restaurar**  y **Cerrar**  del programa en la parte izquierda. Véase Figura 2.3.

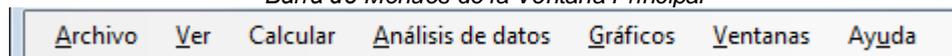
Figura 2.3
ERLA 2011 Estadística ICM/ESPOL
"Barra de título de la ventana principal"



Barra de menues

La barra de menues de ERLA, está ubicada en la parte superior izquierda de la **ventana principal**, debajo de la barra de título y contiene los menues: **Archivo**, **Ver**, **Calcular**, **Análisis de datos**, **Gráficos**, **Ventanas** y **Ayuda**. Véase Figura 2.4.

Figura 2.4
ERLA 2011 Estadística ICM/ESPOL
"Barra de Menues de la Ventana Principal"

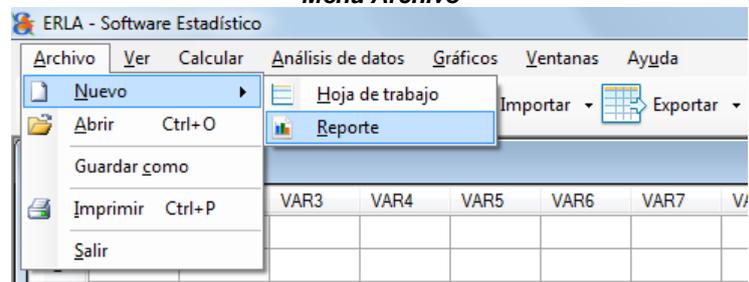


Para desplegar un submenú, basta con pulsar en el ítem contenido por el menú.

Menú Archivo

Al desplegar el menú **Archivo** podrá visualizar los ítems: **Nuevo**, **Abrir**, **Guardar**, **Guardar como**, **Imprimir** y finalmente **Salir**.

Figura 2.5
ERLA 2011 Estadística ICM/ESPOL
"Menú Archivo"



Nuevo: Crea una nueva **ventana de datos** o una nueva **ventana de reporte**.

Abrir: Abre una **ventana de reporte** previamente guardada.

Guardar como: Guarda el contenido de la ventana de reporte que se encuentre activa.

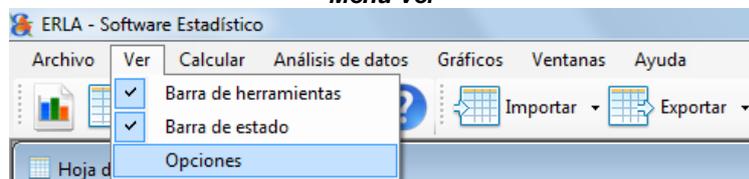
Imprimir: Imprime el contenido de una ventana de reporte.

Salir: Cierra el programa.

Menú Ver

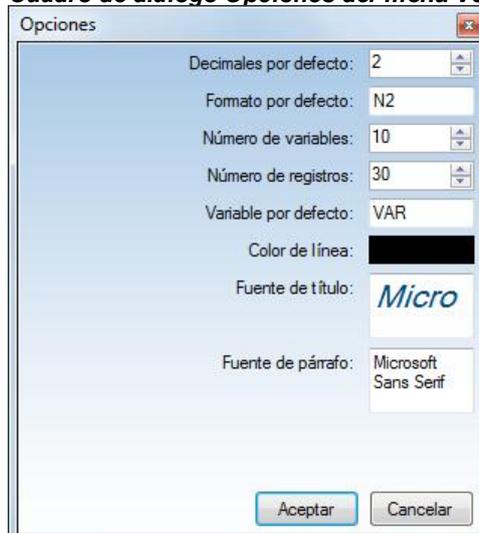
El menú **Ver** permite ocultar o mostrar la barra de herramientas, la barra de estado, y cambiar opciones generales mediante el cuadro de diálogo **Opciones**.

Figura 2.6
ERLA 2011 Estadística ICM/ESPOL
"Menú Ver"



Al pulsar sobre el ítem **Opciones** aparecerá el cuadro de diálogo **Opciones**.

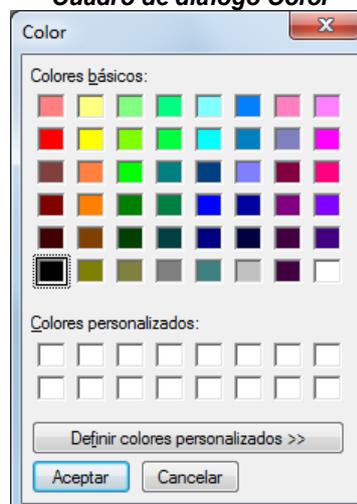
Figura 2.7
ERLA 2011 Estadística ICM/ESPOL
"Cuadro de diálogo Opciones del menú Ver"



Mediante el cuadro de diálogo **Opciones** usted podrá configurar:

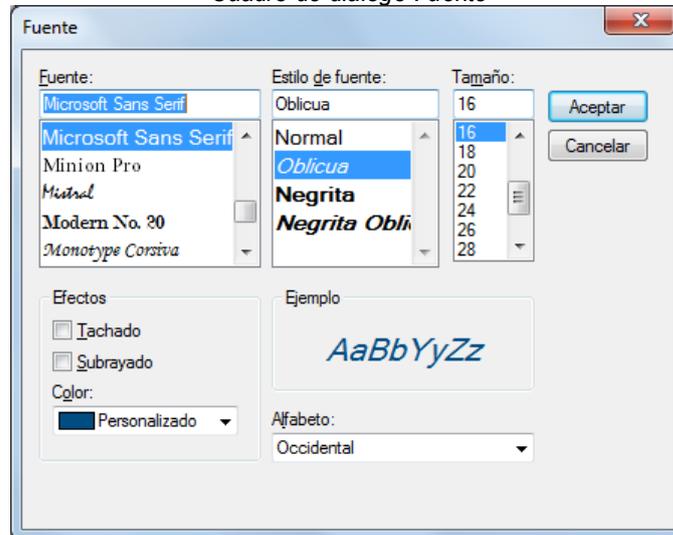
- El número máximo de decimales que desee visualizar en la **ventana de reporte**. Por defecto el software presenta 4 decimales.
- El número de variables que usted desee que aparezcan cada vez que abra una nueva **ventana de datos**. Por defecto aparecerán 10.
- El nombre por defecto que se le asigna a una nueva variable creada. Este nombre se acompaña de cifras enteras que ascienden conforme se crean más variables con el mismo nombre.
- El número de registros que usted desee que aparezcan cada vez que abra una nueva **ventana de datos**. Por defecto aparecerán 10.
- El color de la línea que actúa como separador entre cada resultado generado en la **ventana de reporte**. Véase Figura 2.8.

Figura 2.8
ERLA 2011 Estadística ICM/ESPOL
"Cuadro de diálogo Color"



- Personalizar el tipo de fuente para el título y el párrafo que se presentarán en la **ventana de reporte**. Véase Figura 2.9.

Figura 2.9
ERLA 2011 Estadística ICM/ESPOL
"Cuadro de diálogo Fuente"

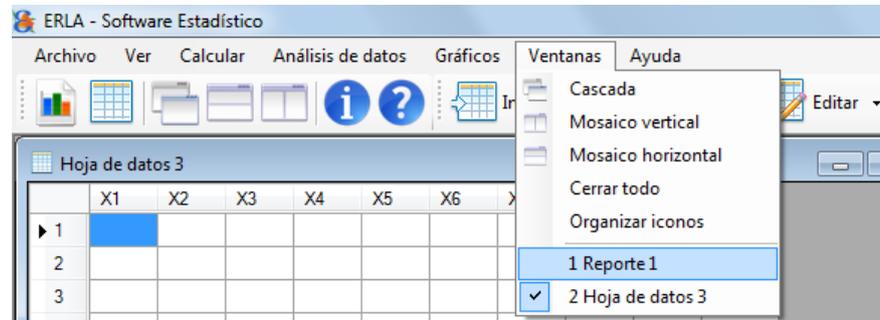


Los menues **Calcular**, **Análisis de datos** y **Gráficos** se presentarán más adelante.

Menú Ventanas

Este menú, permite organizar y cerrar todas las ventanas que estén abiertas en ERLA.

Figura 2.10
ERLA 2011 Estadística ICM/ESPOL
"Menú Ventanas"



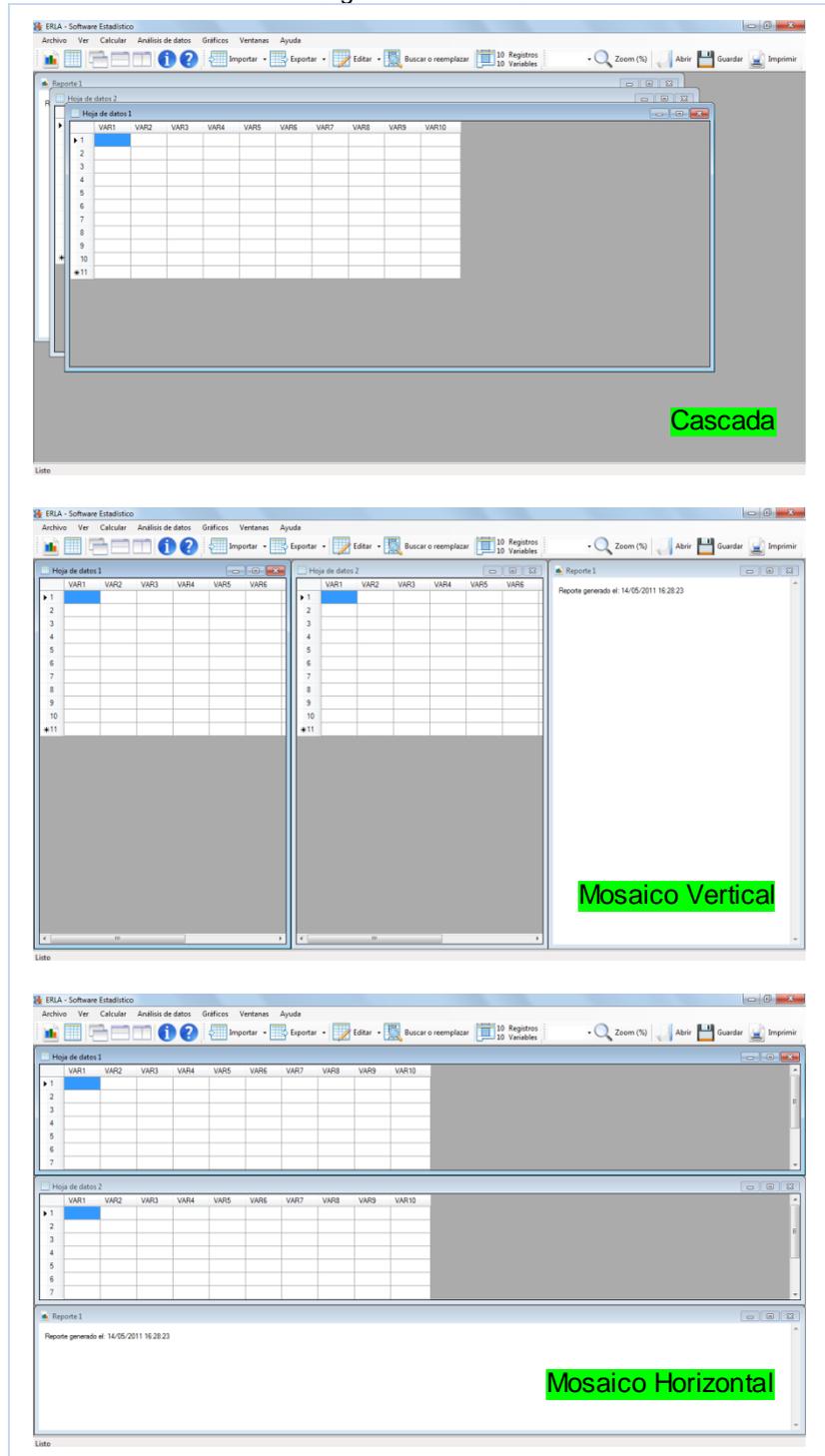
Usted puede activar cualquier **ventana de datos** o **ventana de reporte** cuyo nombre aparezca listado en este menú.

El ítem **Cerrar todo** cierra todas las ventanas que se encuentren dentro de la **ventana principal**.

La opción **Organizar iconos** ordena la disposición de las ventanas minimizadas dentro de la **ventana principal**.

También, se puede organizar las ventanas en **Cascada** , **Mosaico vertical**  y **Mosaico horizontal** . En la Figura 2.11 se presenta gráficamente lo que permite hacer cada una de estas opciones.

Figura 2.11
ERLA 2011 Estadística ICM/ESPOL
"Organizar Ventanas"



Menú Ayuda

Este menú le permite acceder a la información necesaria para ejecutar ERLA, así como datos relacionados con el desarrollo de la versión presente.

Figura 2.12

ERLA 2011 Estadística ICM/ESPOL

“Menú Ayuda”



El ítem **Contenido** muestra el manual de usuario.

El ítem **Acerca de...** muestra información sobre la presente versión del Software y los desarrolladores del mismo.

Barra de herramientas

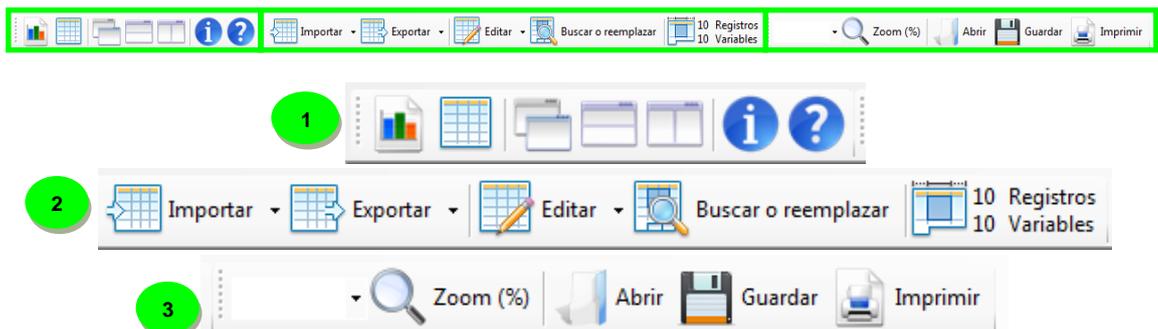
Esta barra está ubicada en la parte superior izquierda de la **ventana principal**, debajo de la **barra de menues**. Está dividida en tres partes denominadas:

1. Barra de herramientas de la **ventana principal**.
2. Barra de herramientas de la **ventana de datos**.
3. Barra de herramientas de la **ventana de reporte**.

Figura 2.13

ERLA 2011 Estadística ICM/ESPOL

“Barra de herramientas de la ventana principal”



Barra de herramientas de la ventana principal

Esta barra de herramientas le permite realizar acciones propias de la **ventana principal**:



Se utiliza para abrir una nueva **ventana de reporte**.



Se utiliza para abrir una nueva **ventana de datos**.



Se utiliza para organizar las ventanas en cascada.



Se utiliza para organizar las ventanas en mosaico horizontal.



Se utiliza para organizar las ventanas en mosaico vertical.



Muestra información sobre la presente versión de ERLA y los desarrolladores del mismo.



Muestra el manual de usuario.

Barra de herramientas de la ventana de datos

Esta barra de herramientas le permitirá realizar acciones sobre la **ventana de datos** en la que se esté trabajando (ventana activa). En la sección correspondiente a la **ventana de datos**, se hablará de cada uno de sus elementos.

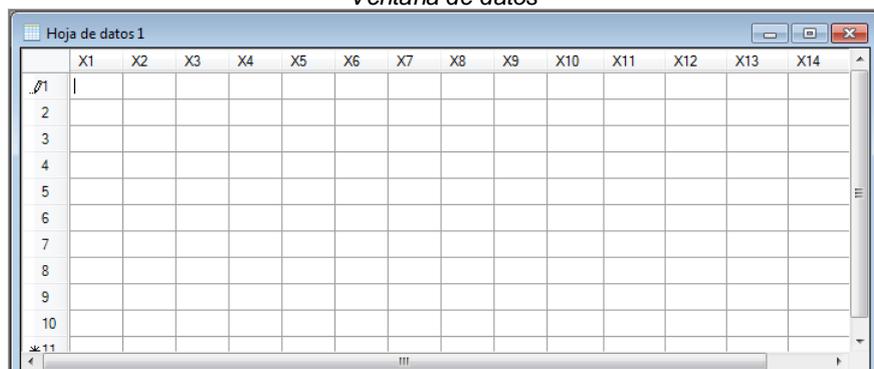
Barra de herramientas de la ventana de reporte

Esta barra de herramientas le permite realizar acciones que afectan exclusivamente a la **ventana de reporte** sobre la cual se esté trabajando (ventana activa). En la sección correspondiente a la **ventana de reporte**, se hablará de cada uno de sus elementos.

Ventana de datos

La **ventana de datos** de ERLA, por lo general está ubicada en la parte izquierda de la **ventana principal**. Es en esta ventana donde se definen las variables que se van a utilizar. Puede crear tantas **ventanas de datos** como desee.

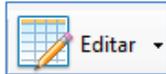
Figura 2.14
ERLA 2011 Estadística ICM/ESPOL
"Ventana de datos"



En la **ventana de datos** usted puede seleccionar celdas, registros, variables o todo el contenido.

Puede realizar distintas acciones en la **ventana de datos** a través de la barra de herramientas de la **ventana de datos** o seleccionando las celdas que desee y pulsando el botón derecho del "ratón".

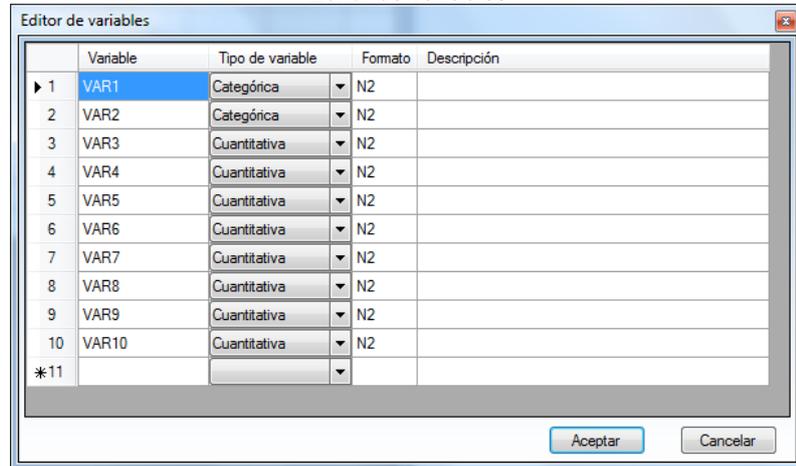
Editor de variables



Usted podrá acceder al **Editor de variables** de ERLA a través de la secuencia:

Barra de herramientas de la ventana de datos ► **Editar** ► **Variables**

Figura 2.15
ERLA 2011 Estadística ICM/ESPOL
"Editor de variables"



El **Editor de variables**, permite agregar nuevas variables, editarlas o eliminarlas. Usted puede definir entre dos tipos de variables: Cuantitativas o categóricas, en el último caso podrá ingresar las etiquetas correspondientes a través de la secuencia:

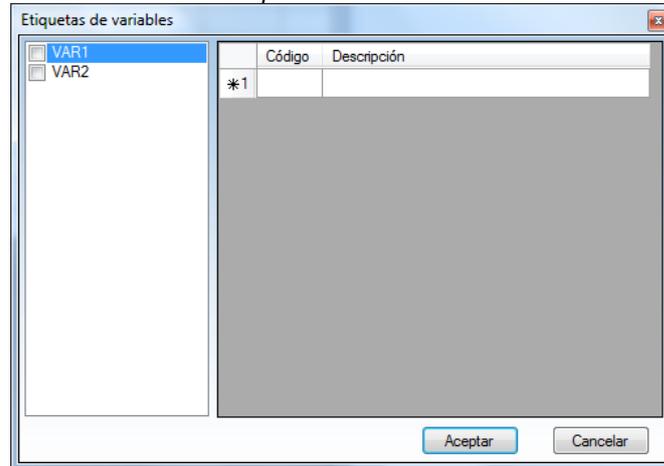
Barra de herramientas de la ventana de datos ► **Editar** ► **Etiquetas**

El cuadro de diálogo **Etiquetas de variables** lista las variables que hayan sido definidas como categóricas en el **Editor de variables**. Para cambiar las etiquetas, seleccione la variable que desee, y en la parte derecha de la ventana escriba el código y la etiqueta correspondiente. Véase Figura 2.16.



Los valores de las variables categóricas se presentan en color rojo oscuro en la **ventana de datos**.

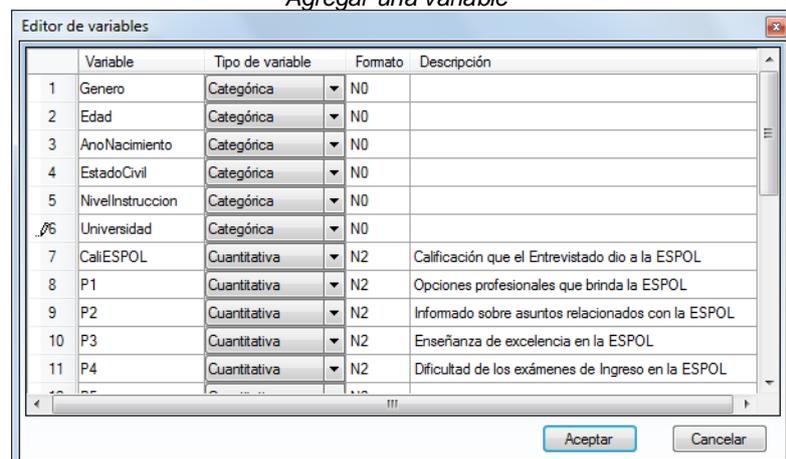
Figura 2.16
ERLA 2011 Estadística ICM/ESPOL
"Etiquetas de variables"



Agregar una variable

1. Vaya a **Barra de herramientas de la ventana de datos** ➤ **Editar** ➤ **Variables**
2. Ubique el cursor en la última fila del cuadro de diálogo **Editor de variables**.
3. Ingrese el nombre de la variable.
4. Seleccione el tipo de variable y el formato.
5. Agregue una descripción de la variable si así lo desea.
6. Pulse el botón **Aceptar**

Figura 2.17
ERLA 2011 Estadística ICM/ESPOL
"Agregar una variable"



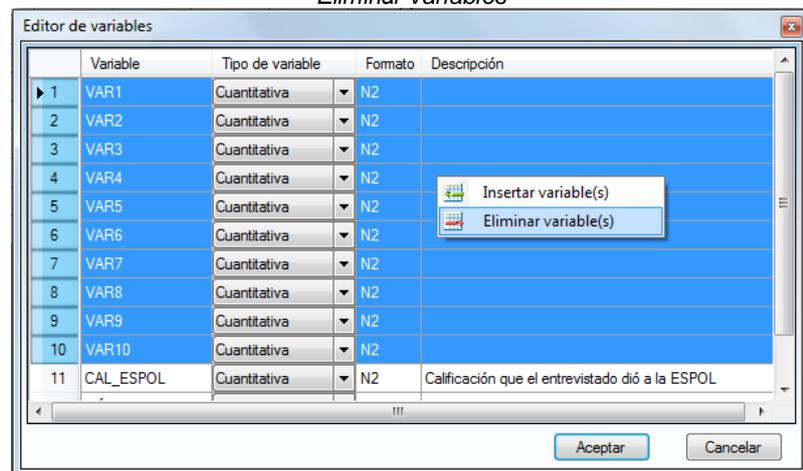


En la columna **Formato** usted indicará el número de decimales con el que desee trabajar colocando la letra N junto al número de decimales (N#).

Eliminar una variable

1. Vaya a **Barra de herramientas de la ventana de datos** ► **Editar** ► **Variables**
2. Seleccione la o las variables que desee eliminar.
3. Pulse el botón derecho del “ratón”.
4. Pulse **Eliminar variable(s)**.
5. Pulse el botón **Aceptar**.

Figura 2.18
ERLA 2011 Estadística ICM/ESPOL
“Eliminar variables”

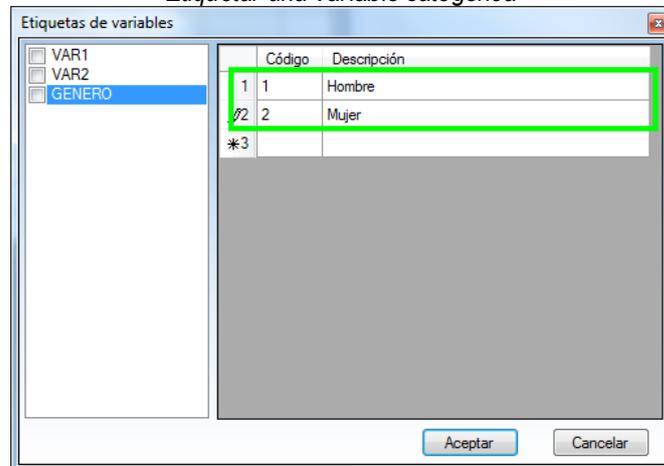


Etiquetar una variable categórica

1. Vaya a **Barra de herramientas de la ventana de datos** ► **Editar** ► **Etiquetas**
2. Seleccione la variable que desee “etiquetar” o “rotular”.
3. Ingrese el código junto con las etiquetas.
4. Si desea repita los pasos 1 y 2 para etiquetar otra variable.

5. Pulse el botón **Aceptar**.

Figura 2.19
ERLA 2011 Estadística ICM/ESPOL
"Etiquetar una variable categórica"



En las variables categóricas se permite todo tipo de caracteres a menos que la variable esté codificada.

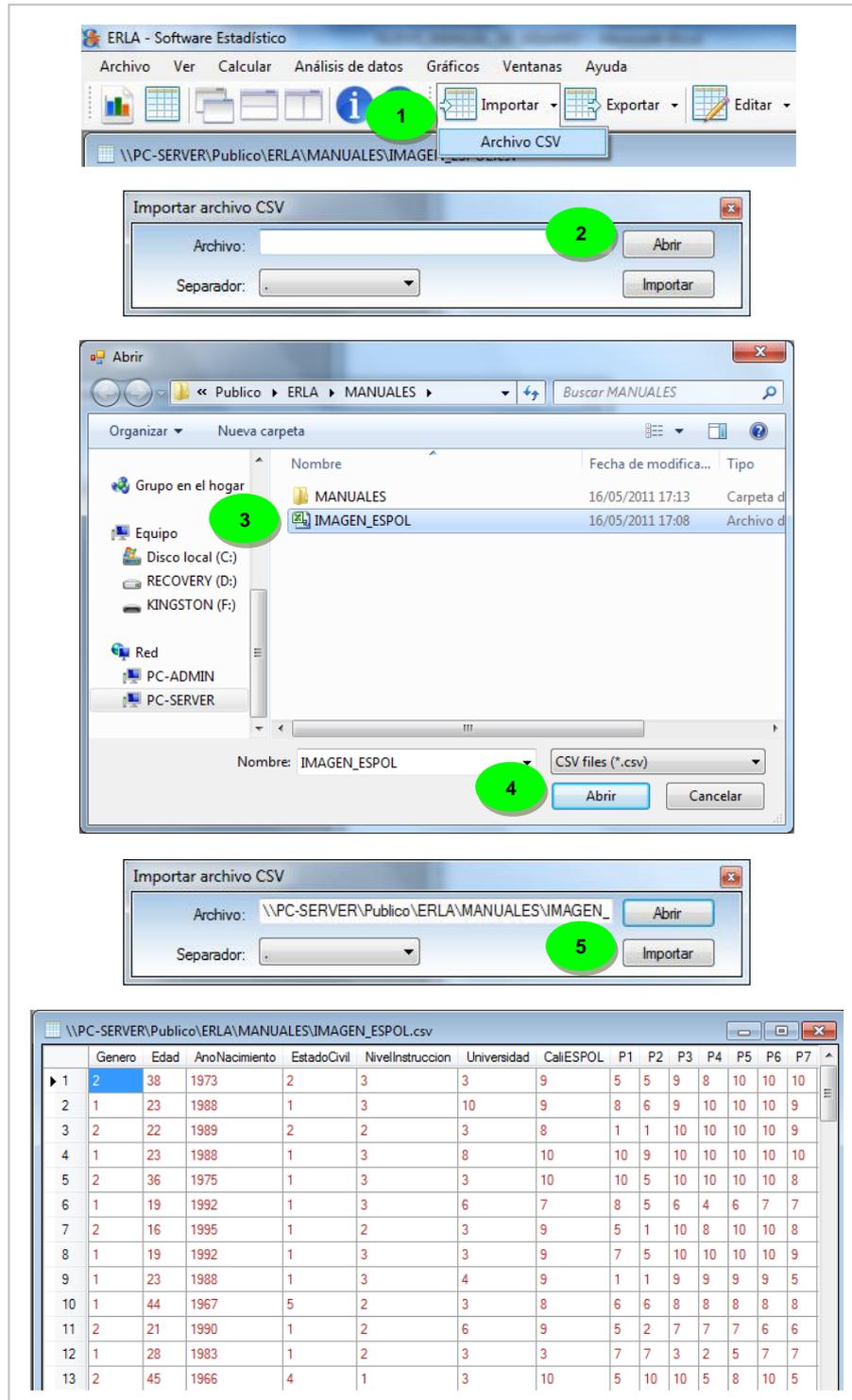
Si marca la casilla de verificación que se encuentra a la izquierda del nombre de la variable, se muestran las etiquetas correspondientes a los códigos de dicha variable en la **ventana de datos**.

Importar datos



1. Vaya a **Barra de herramientas de la ventana de datos** ➤ **Importar** ➤ **Archivo CSV**
2. Pulse el botón **Abrir** del cuadro de diálogo **Importar archivo CSV**.
3. En el cuadro de diálogo **Abrir**, seleccione el archivo csv con los datos.
4. Pulse el botón **Abrir**.
5. Verifique que se presente la ubicación del archivo seleccionado, seleccione el separador y luego pulse en el botón **Importar**

Figura 2.20
ERLA 2011 Estadística ICM/ESPOL
“Importar datos desde un archivo *.csv”





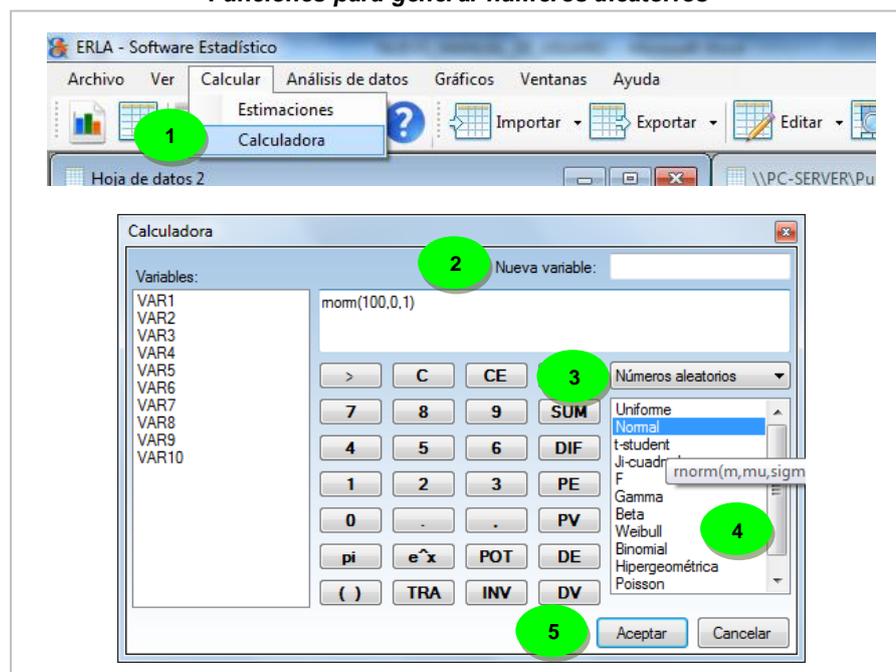
Cuando se importa un archivo, por defecto todas las variables son categóricas.

Generar variables aleatorias

Contando con una **ventana de datos**, realice lo siguiente:

1. Elija **Calcular** ➤ **Calculadora**
2. En el cuadro de diálogo **Calculadora**, ingrese el nombre de la variable (Si lo deja vacío por defecto recibirá el nombre de la operación que vaya a realizar).
3. En la lista desplegable de la derecha seleccione **Números aleatorios**.
4. En el cuadro de lista de abajo seleccione la función de distribución.
5. Ingrese los parámetros de la distribución y pulse el botón **Aceptar**.

Figura 2.21
ERLA 2011 Estadística ICM/ESPOL
“Funciones para generar números aleatorios”



En la Figura 2.21, se van a generar 100 valores aleatorios provenientes de una distribución Normal estándar.



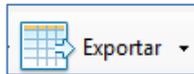
Para obtener información de los parámetros de la distribución que debe ingresar, pase el puntero por el nombre de la distribución.

Con la **Calculadora** usted puede generar variables a partir de otras existentes, ya que esta ofrece una lista de funciones preestablecidas. Véase Tabla 2.1.

Tabla 2.1
ERLA 2011 Estadística ICM/ESPOL
"Lista de funciones predefinidas de la Calculadora"

GRUPO	FUNCIONES	
Funciones fundamentales	Valor Absoluto Raíz Cuadrada Logaritmo de base 10 Logaritmo de base e	Factorial Redondear Ordenar
Números aleatorios	Uniforme Normal T-Student Ji-Cuadrado F-Fisher Gamma	Beta Weibull Binomial Hipergeométrica Poisson Permutación aleatoria
Densidades	Uniforme Normal T-Student Ji-Cuadrado F-Fisher	Gamma Beta Weibull Binomial Hipergeométrica
Distribuciones acumuladas	Normal T-Student Ji-Cuadrado F-Fisher Gamma	Beta Weibull Binomial Hipergeométrica
Funciones inversas	Normal T-Student Ji-Cuadrado F-Fisher Gamma	Beta Weibull Binomial Hipergeométrica
Funciones trascendentes	Seno Coseno Tangente Arcoseno Arcocoseno Arcotangente	Seno hiperbólico Coseno hiperbólico Tangente hiperbólico Arcoseno hiperbólico Arcocoseno hiperbólico Arcotangente hiperbólico

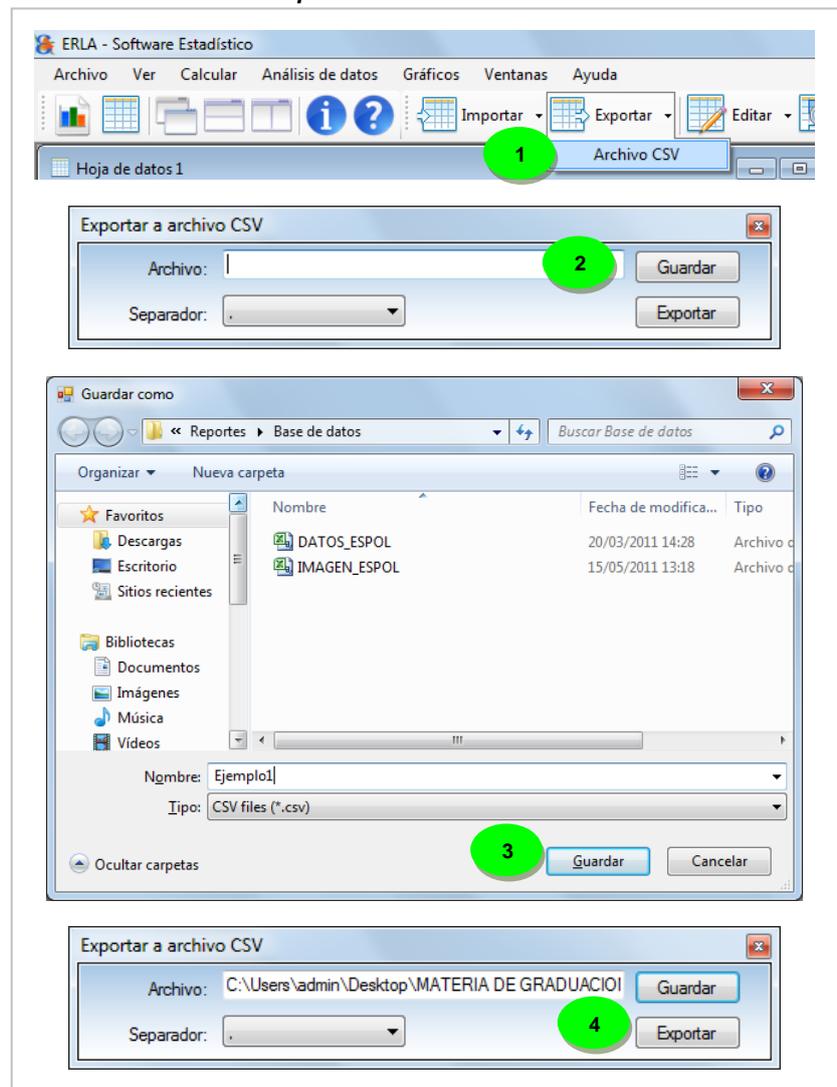
Exportar ventana de datos



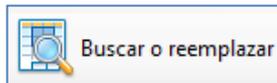
Para exportar los datos de una **ventana de datos**,

1. Vaya a **Barra de herramientas de la ventana de datos** ➤ **Exportar** ➤ **Archivo CSV**,
2. Luego pulse el botón **Guardar**,
3. En el cuadro de diálogo **Guardar como** seleccione la ruta donde desea exportar los datos y escriba el nombre del archivo y pulse en **Guardar**.
4. Finalmente en **Exportar**.

Figura 2.22
ERLA 2011 Estadística ICM/ESPOL
"Exportar ventana de datos"



Buscar y reemplazar

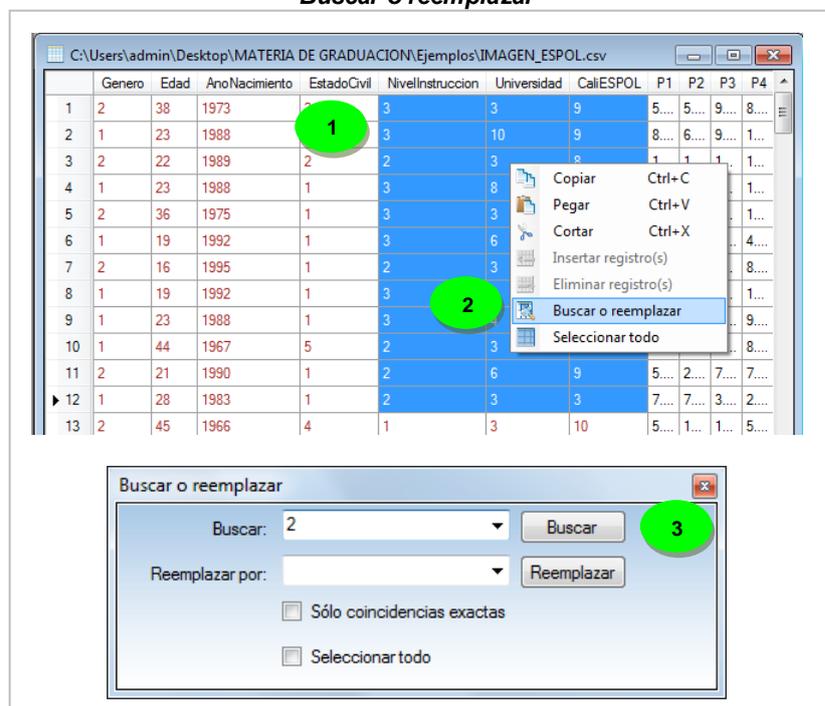


Para buscar y reemplazar algún valor de la **ventana de datos**:

1. Seleccione un conjunto de datos.
2. Pulse en el botón **Buscar o Reemplazar** en la barra de herramientas de la **ventana de datos** o pulse el botón derecho del “ratón” y seleccione **Buscar o Reemplazar**.
3. Ingresar el valor a buscar, luego pulse **Buscar**.

Los resultados o coincidencias se muestran como una selección.

Figura 2.23
ERLA 2011 Estadística ICM/ESPOL
“*Buscar o reemplazar*”

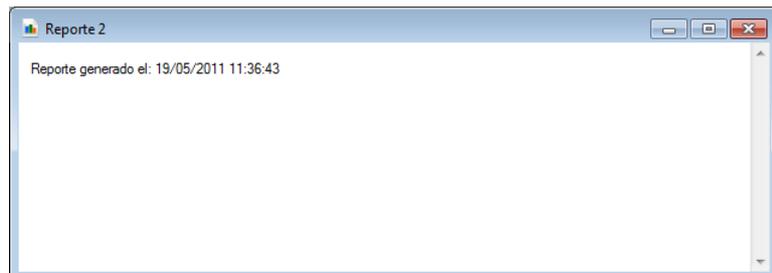


Como puede observar en la Figura 2.23, ERLA también permite reemplazar valores cuyas coincidencias sean exactas o aproximadas, presionando el botón **Reemplazar**.

Ventana de reporte

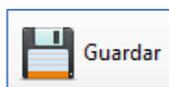
Aquí se muestran cada uno de los resultados que ejecuta ERLA. Por lo general estos resultados se presentan en tablas con sus respectivos títulos.

Figura 2.24
ERLA 2011 Estadística ICM/ESPOL
“Ventana de reporte de ERLA”



El usuario puede ampliar el tamaño de los resultados del informe mediante la lista **Zoom (%)** de la **Barra de herramientas de la ventana de reporte**.

Guardar una ventana de reporte

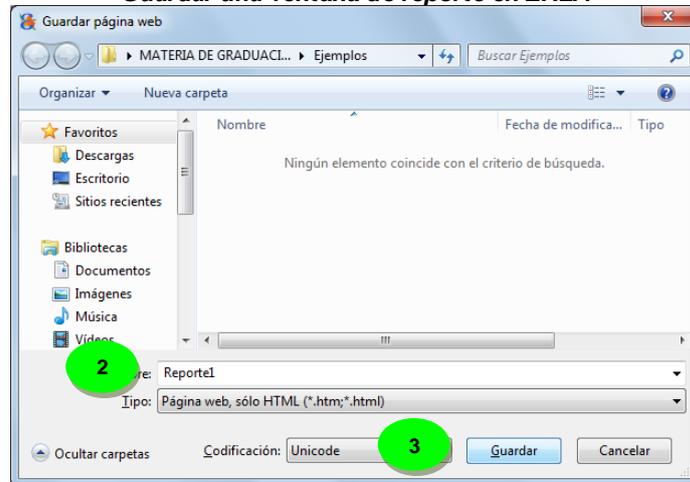


Para guardar una **ventana de reporte**:

1. **Vaya a la Barra de herramientas de la ventana de reporte**
➤ Guardar
2. **En Nombre escriba el nombre del archivo, por ejemplo *Reporte1*, y en Tipo elija Página Web, sólo HTML. Es importante señalar que los reportes son generados en lenguaje HTML, por lo que sus estilos pueden ser personalizados.**
3. **Pulse el botón Guardar**

Véase Figura 2.25

Figura 2.25
ERLA 2011 Estadística ICM/ESPOL
“Guardar una ventana de reporte en ERLA”



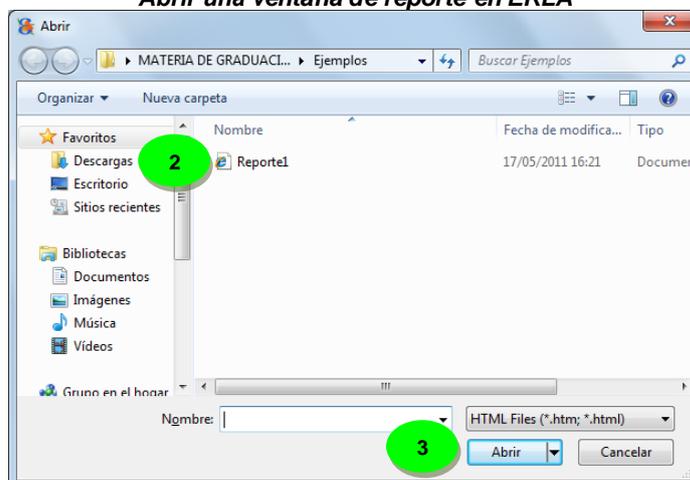
Abrir una ventana de reporte



Para abrir una **ventana de reporte**:

1. **Vaya a la** Barra de herramientas de la ventana de reporte
➤ **Abrir**
2. **Seleccione el archivo.**
3. **Pulse el botón Abrir.**

Figura 2.26
ERLA 2011 Estadística ICM/ESPOL
“Abrir una ventana de reporte en ERLA”



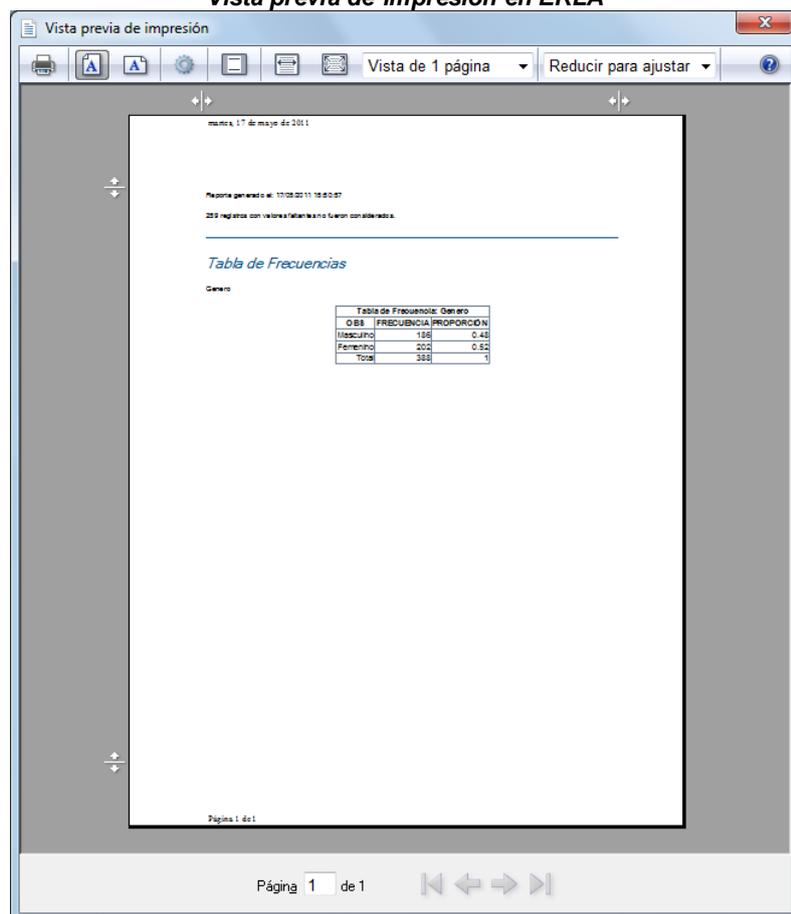
Imprimir reporte



Para imprimir un reporte, pulse en **Imprimir** en la **barra de herramientas de la ventana de reporte**. Véase Figura 2.27.

La ventana **Vista previa de impresión**, está formada por una **barra de título**, una **barra de herramientas** y un **pie de página**. Al hacer contacto con cada uno de los botones de la barra de herramientas, usted puede visualizar un mensaje indicando lo que éste hace.

Figura 2.27
ERLA 2011 Estadística ICM/ESPOL
“Vista previa de impresión en ERLA”



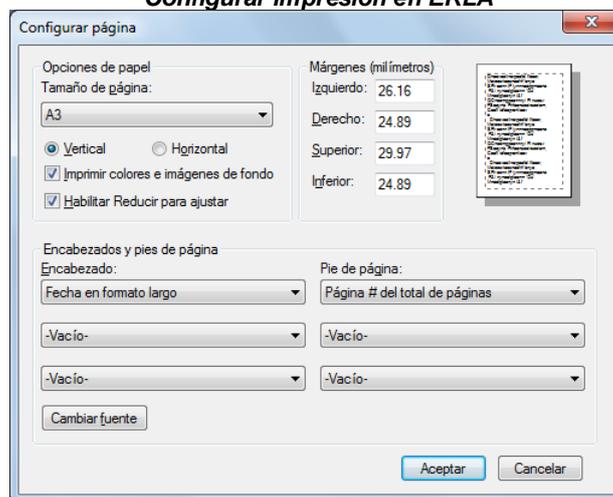
A través de la Barra de herramientas de la ventana “Vista previa de Impresión” usted podrá:

1. Seleccionar la impresora para ejecutar la impresión del reporte.

2.  Cambiar la orientación de la página.

3.  Configurar la página. Véase Figura 2.28.

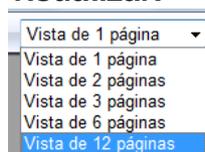
Figura 2.28
ERLA 2011 Estadística ICM/ESPOL
"Configurar impresión en ERLA"



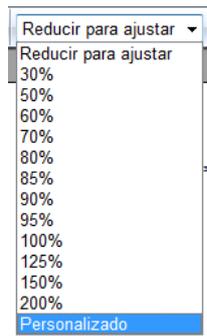
4.  Mostrar y ocultar encabezados y pié de página.

5.  Ver ancho total de la página o verla completa.

6. Seleccionar el número de páginas que desea visualizar.



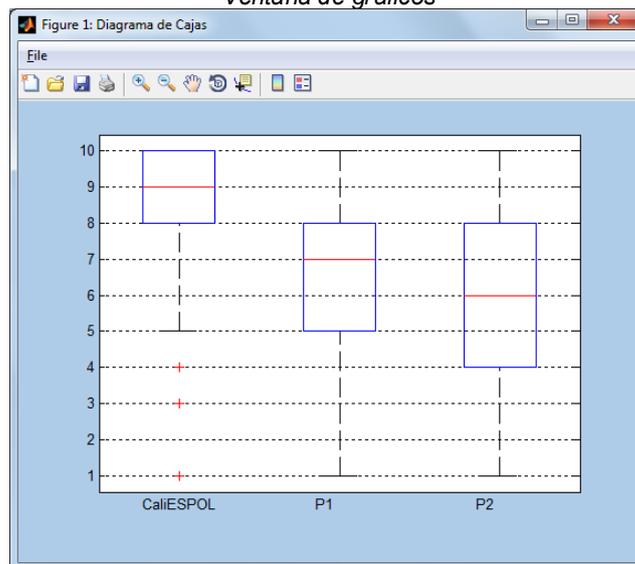
7. Ajustar el contenido de la página.



Ventana de gráficos

A diferencia de las dos ventanas anteriores, las cuales aparecen al abrir ERLA, la **ventana de gráficos** solo aparecerá cuando se llame alguno de ellos. Esta ventana está formada por una **barra de título**, una **barra de menú** y una **barra de herramientas** las cuales son propias de MATLAB.

Figura 2.29
ERLA 2011 Estadística ICM/ESPOL
"Ventana de gráficos"



En la Figura 2.29 se muestra la **barra de herramientas de la ventana de gráficos**, y el menú **File** a través del cual el usuario puede:

1.  Abrir una nueva ventana de gráfico.
2.  Abrir un gráfico con la extensión **.fig* previamente guardado.
3.  Guardar el gráfico con la extensión **.fig* / **.jpg* / **.png*
4.  Imprimir el gráfico. Véase Figura 2.30.

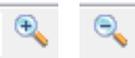
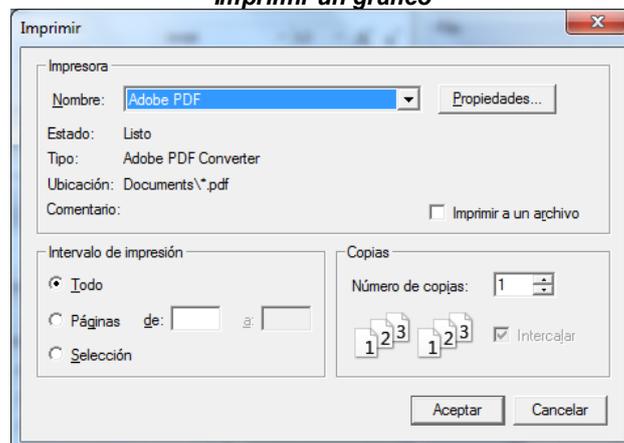
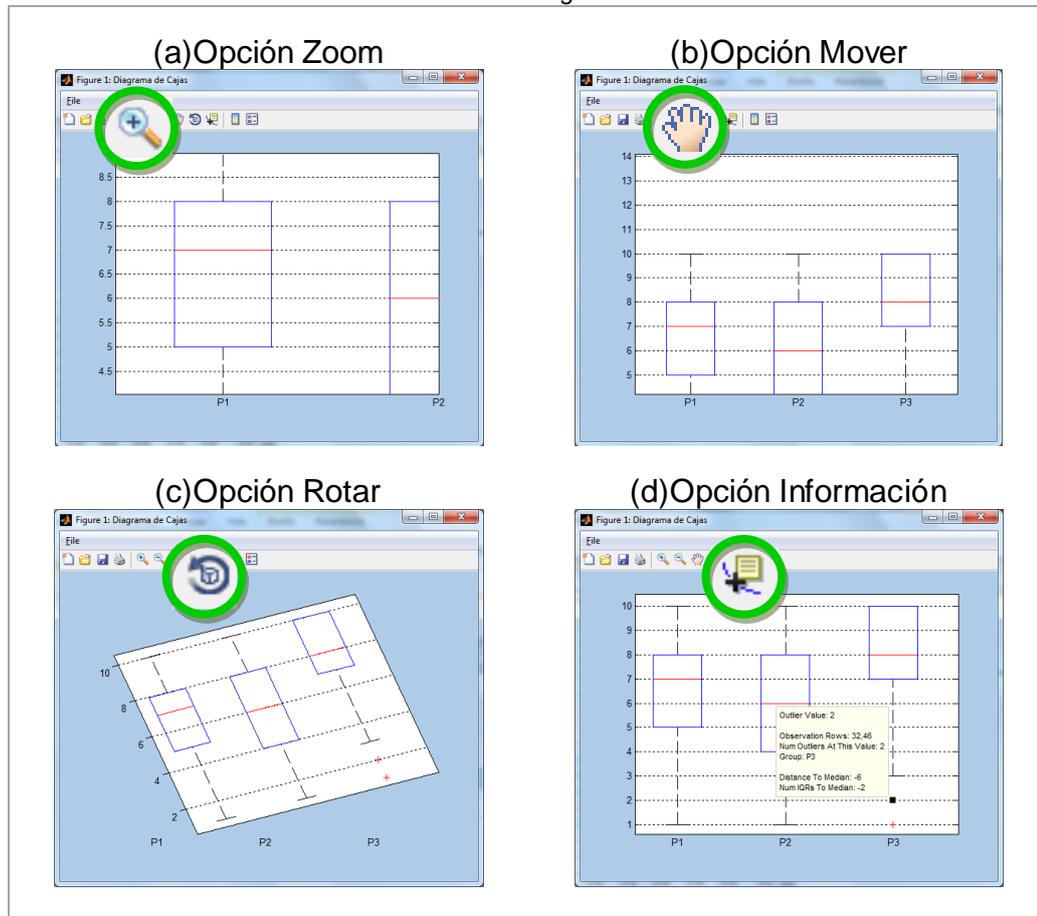
5.  Hacer un zoom al gráfico ampliando o reduciendo las escalas de cada eje.
6.  Moverse a través del gráfico.
7.  Rotar el gráfico.
8.  Obtener información del gráfico.
9.  Además, se puede visualizar una barra de colores que indican los valores más altos y más bajos de la función graficada a través de una escala de colores. Ésta opción es más usada en gráficos 3D.

Figura 2.30
ERLA 2011 Estadística ICM/ESPOL
"Imprimir un gráfico"



Adicionalmente podrá exportar el gráfico y abrir la vista previa de impresión a través del menú File. Véase Figura 2.31.

Figura 2.31
ERLA 2011 Estadística ICM/ESPOL
"Modificar un gráfico"



Barra de estado

La **barra de estado** se encuentra ubicada en la parte inferior izquierda de la **ventana principal** y permite saber si un proceso se ha completado exitosamente o no.

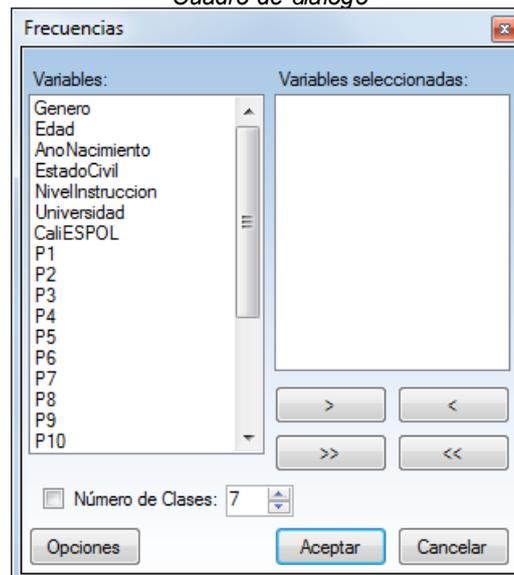
Figura 2.32
ERLA 2011 Estadística ICM/ESPOL
"Barra de estado"

Listo

Cuadro de diálogo

Un **cuadro de diálogo** es un tipo especial de ventana que solicita alguna clase de información al usuario. Los cuadros de diálogo en ERLA tienen un formato determinado. Véase Figura 2.34.

Figura 2.33
ERLA 2011 Estadística ICM/ESPOL
"Cuadro de diálogo"



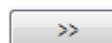
Por lo general, a la derecha de un cuadro de diálogo se listan las variables que se encuentren en la **ventana de datos**. A la izquierda de la ventana, usted deberá seleccionar las variables necesarias para el análisis. Esto lo puede realizar a través de los botones:



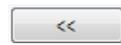
Agregar una variable a la lista de variables seleccionadas.



Quitar una variable de la lista de variables seleccionadas.



Agregar todas las variables a la lista de variables seleccionadas.



Quitar todas las variables de la lista de variables seleccionadas.

Por otra parte, el botón **Opciones** muestra un nuevo cuadro de diálogo que le permitirá personalizar los resultados.



Cada vez que realice operaciones estadísticas a partir de los menús **Análisis de datos** o **Gráficos**, se abre un cuadro de diálogo en el cual debe seleccionar las variables a emplear según el tipo de análisis o gráfico a presentar.



ESTADÍSTICA DESCRIPTIVA

Objetivos

- Obtener estadísticas descriptivas y tablas de frecuencias.
- Elaborar gráficos estadísticos.

Revisión general

En este capítulo se proporciona la información necesaria para que pueda obtener las estadísticas básicas de las variables objeto de estudio, estos son los estadísticos descriptivos más usuales, así como tablas de frecuencias, diagrama de barras, histogramas de frecuencia, diagramas de cajas, y gráficos de dispersión.

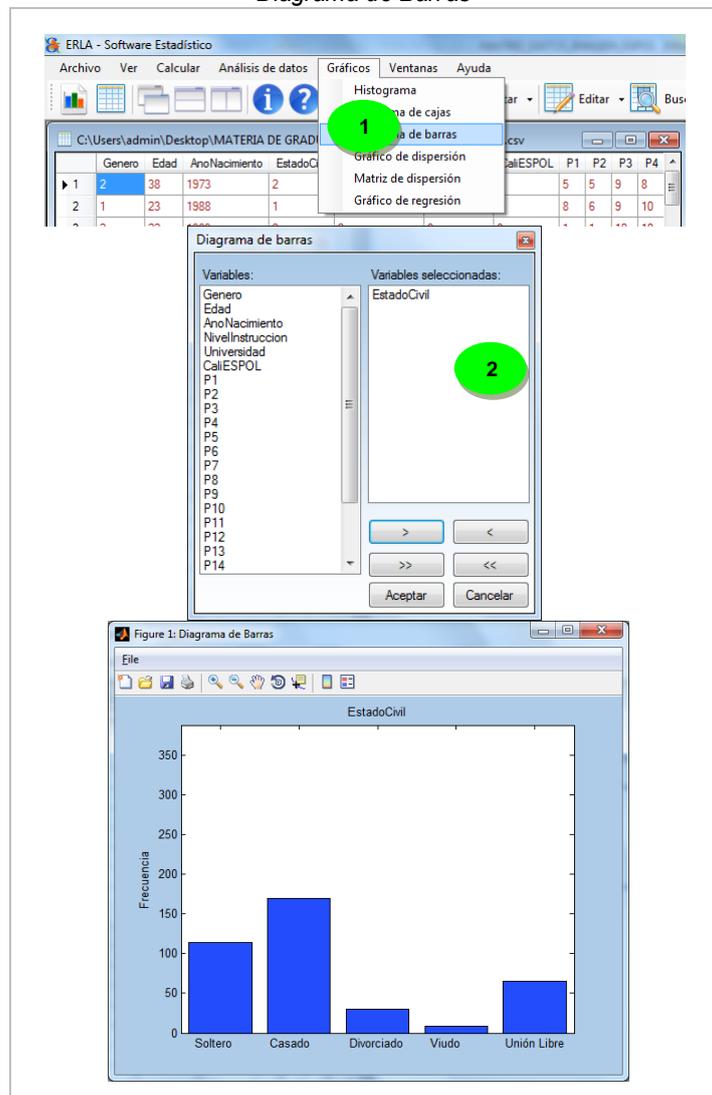
Gráficos

Diagrama de barras

Para realizar un diagrama de barras en ERLA siga la secuencia:

1. **Barra de menú**►**Gráficos**►**Diagrama de barras**
2. En el cuadro de diálogo **Diagrama de barras** seleccione las variables que desee y luego pulse el botón **Aceptar**. Se abren tantas **ventanas de gráficos** como variables seleccionadas. Véase Figura 3.2.

Figura 3.1
ERLA 2011 Estadística ICM/ESPOL
"Diagrama de Barras"



Histograma

Para realizar un histograma en ERLA siga la secuencia:

1. **Barra de menú**►**Gráficos**►**Histograma**
2. En el cuadro de diálogo **Histograma** seleccione las variables que desee, el número de clases y luego pulse el botón **Aceptar**. Aparecerán tantas **Ventanas de Gráficos** como variables haya seleccionado. Véase Figura 3.2.

Figura 3.2
ERLA 2011 Estadística ICM/ESPOL
"Histograma"

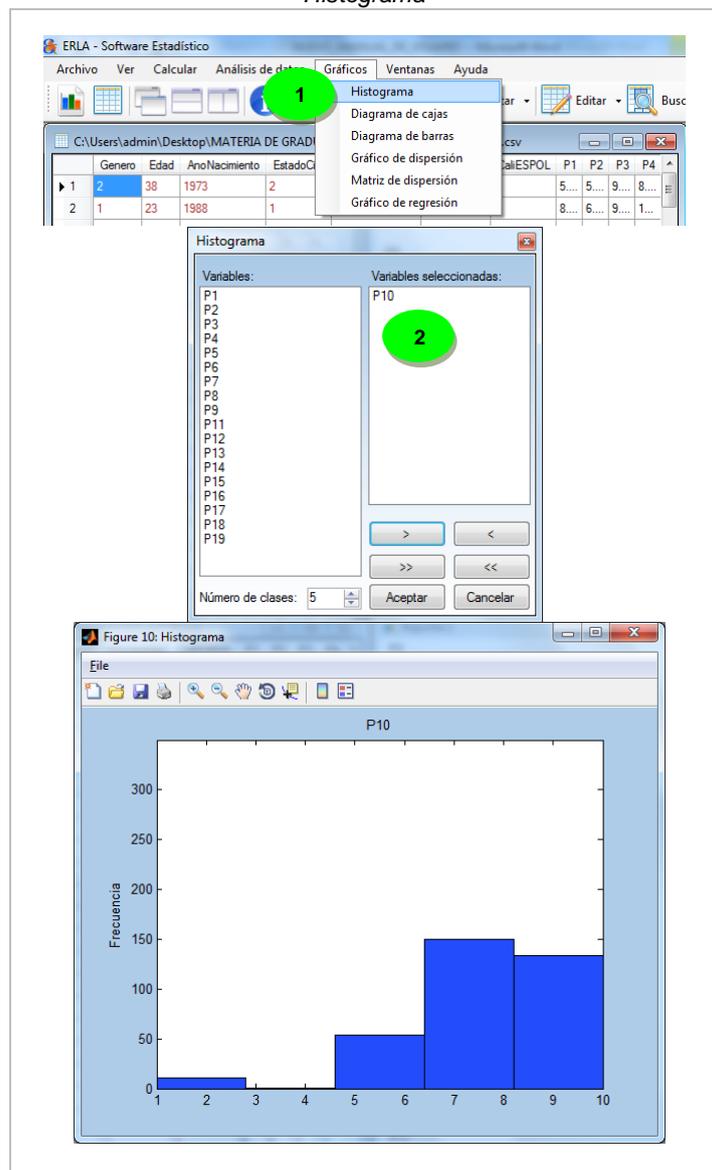


Diagrama de cajas

Para realizar un diagrama de cajas en ERLA siga la secuencia:

1. **Barra de menú>Gráficos>Diagrama de cajas**
2. En el cuadro de diálogo **Diagrama de cajas** seleccione las variables que desee y luego pulse el botón **Aceptar**. Se abrirán tantas **ventanas de gráficos** como variables seleccionadas. La opción **Todo en uno** muestra todos los gráficos en una misma ventana. Véase Figura 3.3.

Figura 3.3
ERLA 2011 Estadística ICM/ESPOL
"Diagrama de cajas"

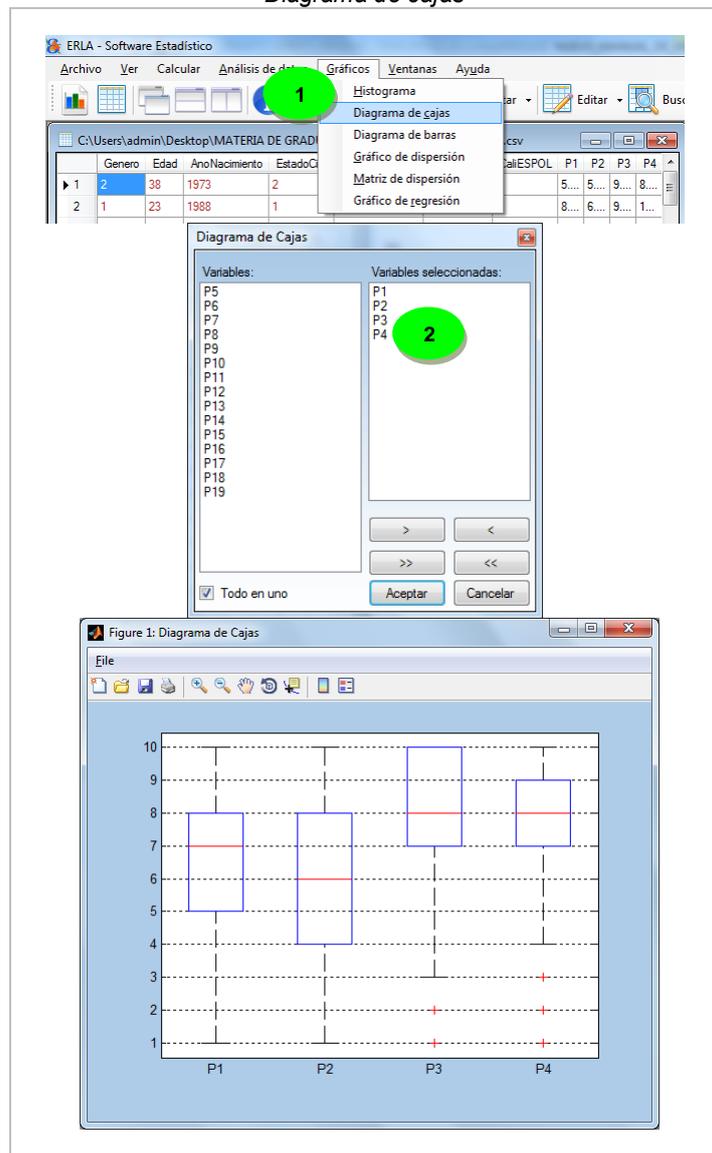
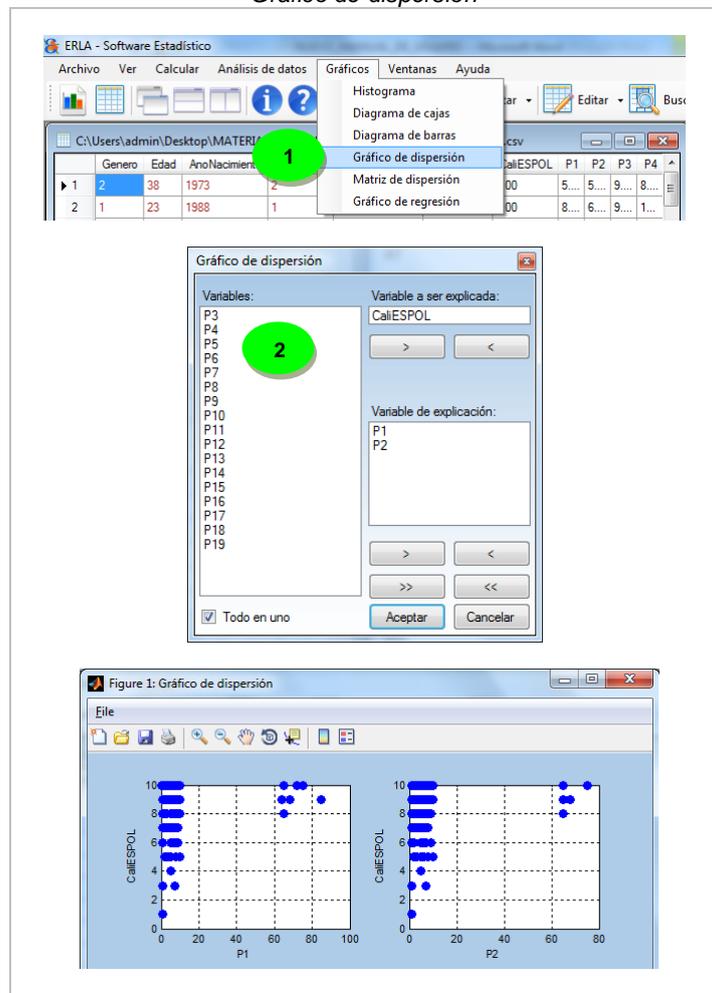


Gráfico de dispersión

Para realizar un gráfico de dispersión en ERLA siga la secuencia:

1. **Barra de menú** > **Gráficos** > **Gráfico de dispersión**
2. En el cuadro de diálogo **Gráfico de dispersión** seleccione las variables que desee y luego pulse el botón **Aceptar**. Aparecerán tantas **ventanas de gráficos** como variables haya seleccionado al menos que seleccione la opción **Todo en uno**, para mostrar todos los gráficos en una misma ventana. Véase Figura 3.4.

Figura 3.4
ERLA 2011 Estadística ICM/ESPOL
"Gráfico de dispersión"



Matriz de dispersión

Para realizar una matriz de dispersión en ERLA siga la secuencia:

1. **Barra de Menú**►**Gráficos**►**Matriz de dispersión**
2. En el cuadro de diálogo **Matriz de dispersión** seleccione las variables que desee y luego pulse el botón **Aceptar**. Véase Figura 3.5.

Figura 3.5
ERLA 2011 Estadística ICM/ESPOL
"Matriz de dispersión"

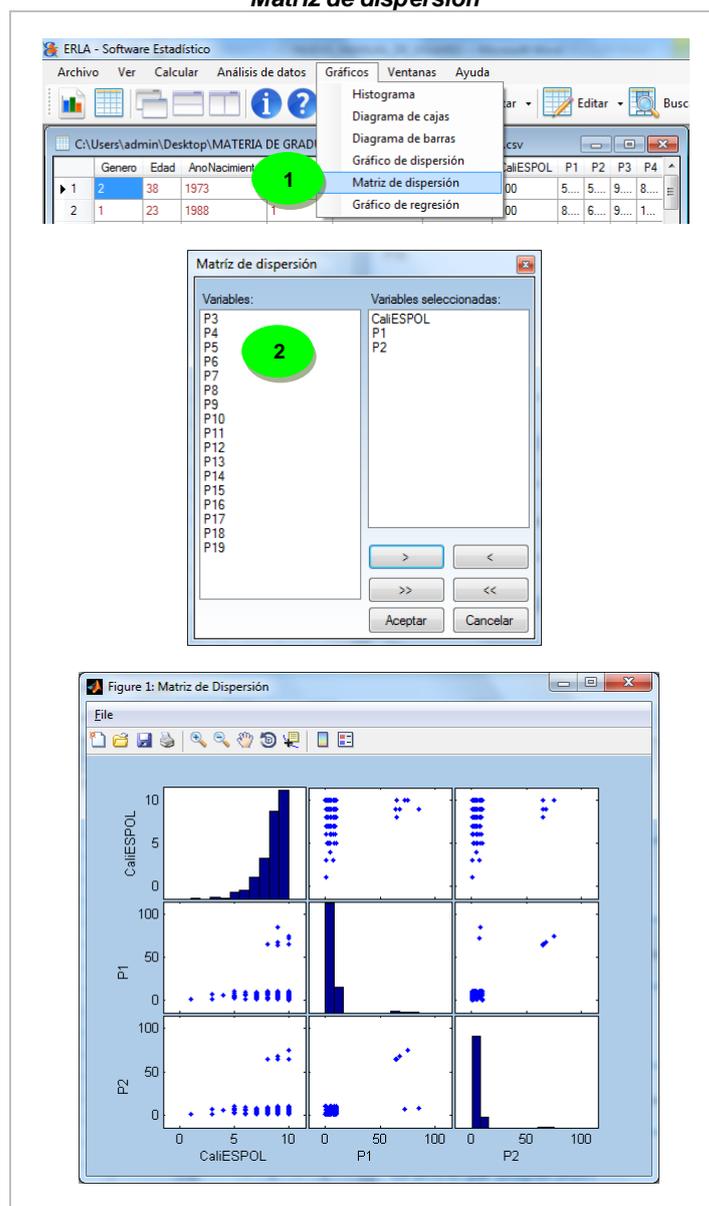


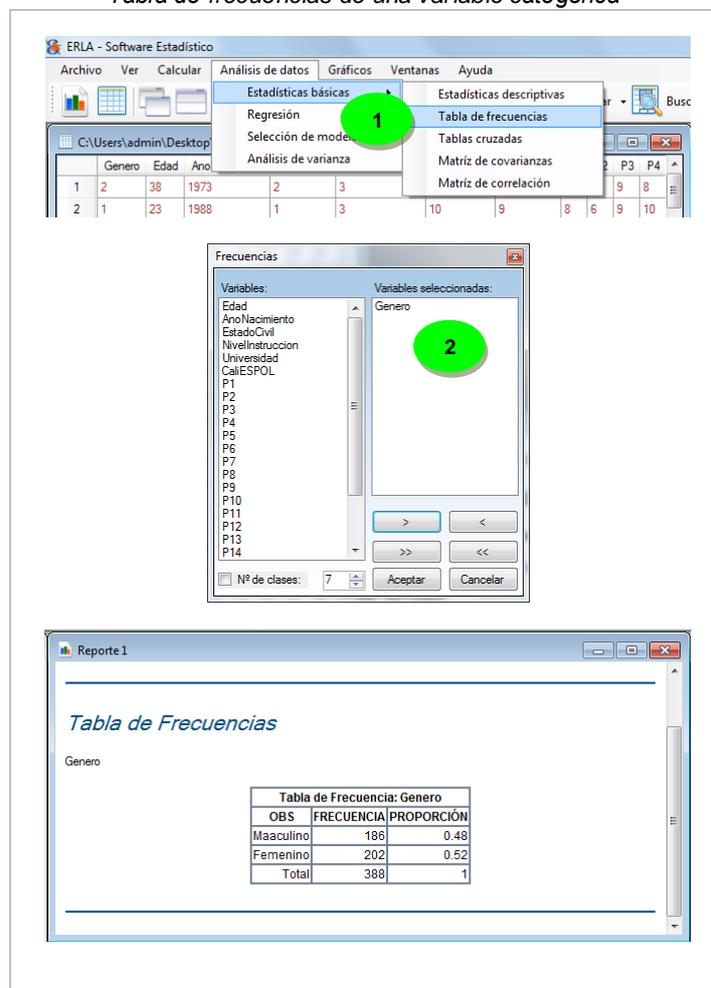
Tabla de frecuencia

Variable categórica

Para obtener una tabla de frecuencias de una variable categórica en ERLA, siga la secuencia:

1. **Barra de menues** ➤ **Análisis de datos** ➤ **Estadísticas básicas** ➤ **Tabla de frecuencias**
2. En el cuadro de diálogo **Frecuencias**, debe seleccionar las variables. Por ser variable categórica la casilla **Número de clases** debe estar desmarcada. Pulse el botón **Aceptar**. Véase Figura 3.6.

Figura 3.6
ERLA 2011 Estadística ICM/ESPOL
"Tabla de frecuencias de una variable categórica"

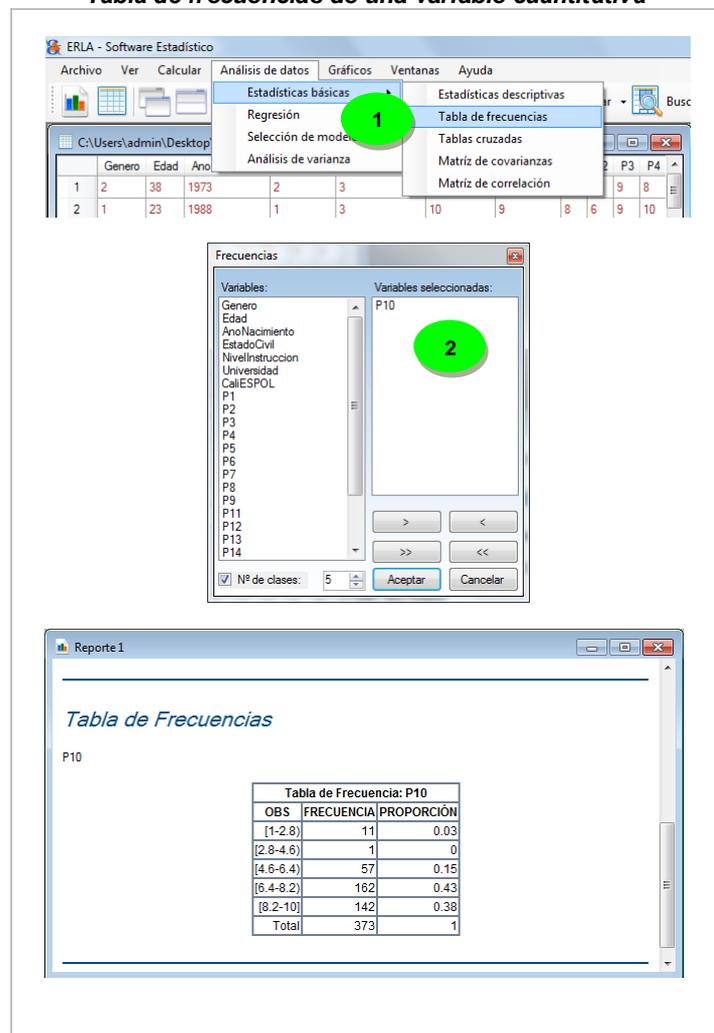


Variable cuantitativa

Para obtener una tabla de frecuencias de una variable cuantitativa en ERLA, siga la secuencia:

1. **Barra de menues** ➤ **Análisis de datos** ➤ **Estadísticas básicas** ➤ **Tabla de frecuencias**
2. En el cuadro de diálogo **Frecuencias**, debe seleccionar las variables. Verifique que la casilla **Número de clases** este marcada y coloque el número de clases que desee y pulse el botón **Aceptar**. Véase Figura 3.7.

Figura 3.7
ERLA 2011 Estadística ICM/ESPOL
"Tabla de frecuencias de una variable cuantitativa"



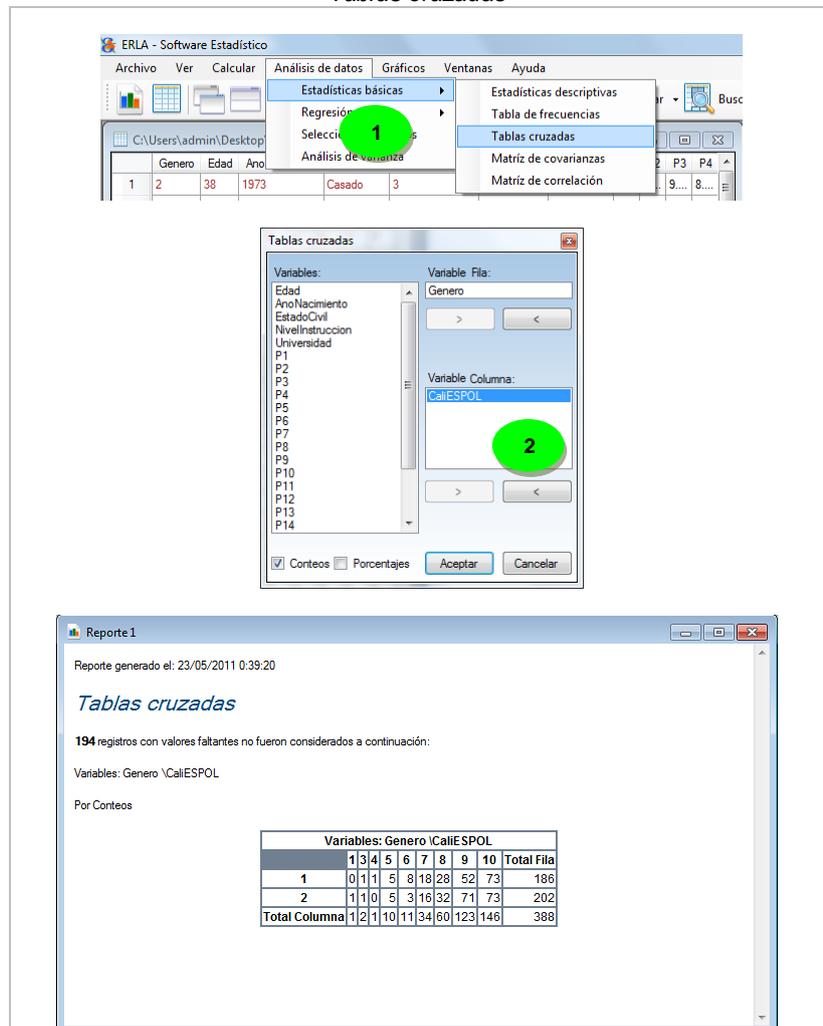
Tablas cruzadas

Para crear tablas cruzadas se debe seguir la secuencia:

Barra de menú ➤ **Análisis de datos** ➤ **Estadísticas básicas** ➤ **Tablas cruzadas**

Seleccione las variables que desee. Además, puede seleccionar las opciones **Conteos** o **Porcentajes**, si quiere que los resultados se presenten como conteo de casos o porcentaje de casos respectivamente. Véase Figura 3.8.

Figura 3.8
ERLA 2011 Estadística ICM/ESPOL
"Tablas cruzadas"



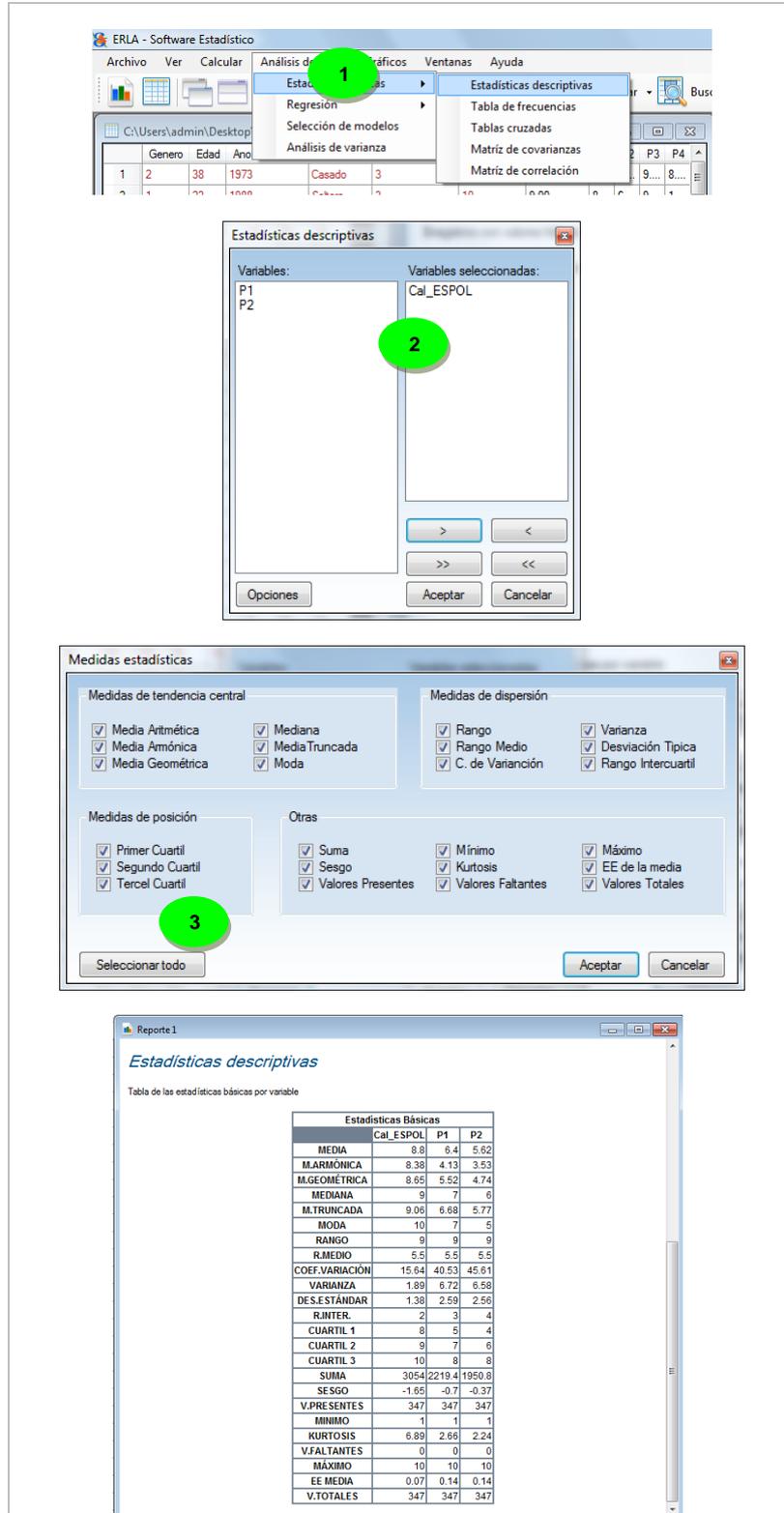
Estadísticas descriptivas

En esta parte el usuario obtiene las estadísticas básicas de las variables de interés, estas son: Media aritmética, media armónica, media geométrica, mediana, media truncada, moda, rango, rango medio, coeficiente de variación, varianza, desviación típica, rango intercuartil, cuartiles, suma, sesgo, mínimo, kurtosis, valores faltantes, máximo, error estándar de la media y valores totales.

Para calcular las estadísticas descriptivas siga la secuencia:

1. **Barra de menú** > **Análisis de datos** > **Estadísticas básicas** > **Estadísticas descriptivas**.
2. En el cuadro de diálogo **Estadísticas descriptivas** seleccione las variables y luego pulse el botón **Opciones**.
3. En el cuadro de diálogo **Medidas estadísticas** se puede seleccionar los estadísticos que se desee mostrar. Dentro de esta ventana, el botón **Seleccionar todo** marca y desmarca todas las opciones de medidas estadísticas disponibles. Véase Figura 3.9.

Figura 3.9
ERLA 2011 Estadística ICM/ESPOL
“Estadísticas descriptivas”





CORRELACIÓN, REGRESIÓN Y ANÁLISIS DE VARIANZA

Objetivos

- Elaborar modelos de: Regresión Lineal, Poisson, Logística, Ridge y Robusta.
- Obtener gráficos e indicadores que permitan evaluar la calidad del modelo de regresión lineal.
- Seleccionar el mejor modelo considerando todas las variables de explicación posibles.
- Obtener la correlación entre variables.
- Ejecutar el Análisis de Varianza (ANOVA) de una y dos vías.

Revisión general

En este capítulo se presenta paso a paso como elaborar modelos de regresión con la ayuda de ERLA.

Para el caso de la regresión lineal se usa el caso de estudio planteado en el capítulo 1 esto es “Imagen de la ESPOL en Guayaquil”, y se ilustra mediante un ejemplo como obtener modelos de regresión lineal simple y múltiple considerando términos polinómicos e interacciones.

Para otros tipos de regresión y análisis de varianza se utilizan datos provenientes de otras fuentes.

Para este capítulo, se utilizarán las siguientes variables del caso de estudio “Imagen de la ESPOL en Guayaquil”.

ERLA	DESCRIPCIÓN
Genero	-
Edad	-
AnoNacimiento	Año de Nacimiento
EstadoCivil	Estado Civil
NivelInstruccion	Nivel de Instrucción
Universidad	Universidad
CaliESPOL	Calificación que el estudiante dio a la ESPOL
P1	Opciones profesionales que brinda la ESPOL.
P2	Asuntos relacionados con la ESPOL.
P3	Excelencia de la enseñanza en la ESPOL.
P4	Dificultad de los exámenes de ingreso.
P5	Seriedad con la que se examina a los estudiantes.
P6	Contenido en Ciencias.
P7	Contenido en Informática.
P8	Contenido Técnico.
P9	Responsabilidad de los estudiantes.
P10	Honestidad de los estudiantes.
P11	Formar Profesionales de excelencia.
P12	Desempeño a nivel Profesional.
P13	Ventajas competitivas
P14	Ventajas frente a competidores de otras universidades.
P15	Impacto de la ESPOL en la Comunidad.
P16	Investigación Científica.
P17	Programas Culturales.
P18	Preocupación por el medio Ambiente.
P19	Relación con Partido Político alguno.

Las variables P1 a P19 son proposiciones calificadas en una escala del 1 al 10, donde 1 indica “Total desacuerdo” con lo planteado, y 10 “Total acuerdo”.

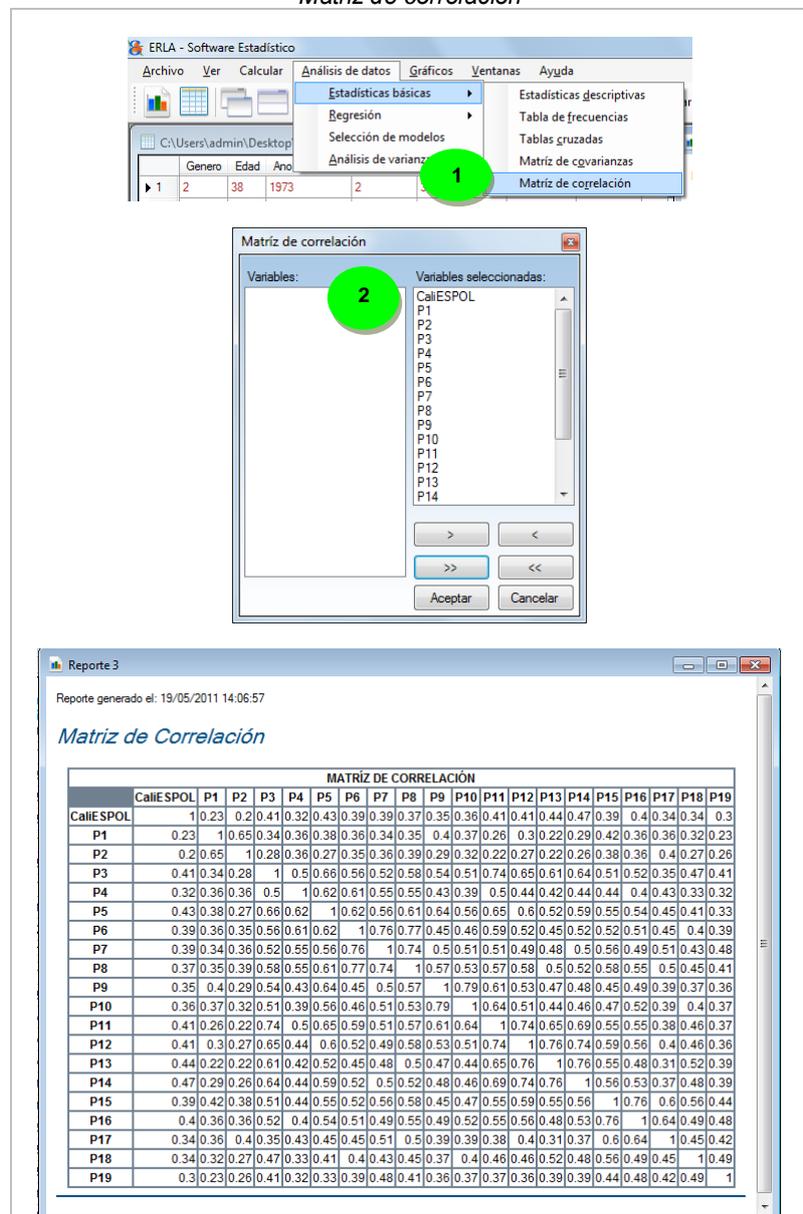
Correlación

Suponiendo que ya se han obtenido las estadísticas descriptivas, tabulaciones y gráficos de las variables de interés, lo siguiente es obtener las correlaciones entre cada par de ellas. Para ello siga la secuencia:

- 1. Barra de menús > Análisis de datos > Estadísticas básicas > Matriz de correlación**

2. Seleccione las variables que desee y luego pulse el botón **Aceptar**.

Figura 4.1
ERLA 2011 Estadística ICM/ESPOL
"Matriz de correlación"



Interpretación de resultados

En la matriz de correlación, se puede observar qué pares de variables son las que están altamente correlacionadas. Esto nos puede dar una idea de que variables utilizar para elaborar el modelo de regresión.

Regresión

Regresión lineal simple

En la matriz de correlación se pudo observar que las proposiciones *P9* (*Identifico a los estudiantes de la ESPOL por su responsabilidad*) y *P10* (*Identifico a los estudiantes de la ESPOL por su honestidad*) están altamente correlacionadas.

Veamos que sucede si se plantea un modelo de regresión lineal con este par de variables. Proceda de la siguiente forma:

1. **Barra de menús > Análisis de datos > Regresión > Regresión lineal**
2. Seleccione la variable a ser explicada y la variable de explicación.
3. Pulse el botón **Aceptar**.

En la Figura 4.2 se pueden observar los resultados obtenidos.

Regresión lineal simple polinómica

Para obtener un modelo de regresión lineal simple polinómico en ERLA, es necesario utilizar la Calculadora; para ello vaya a:

1. **Barra de menús > Calcular > Calculadora.**
2. Realice la operación necesaria, en este caso $P9.^2$ y deje el cuadro de texto **Nombre de la variable** en blanco para que tome el nombre de la operación.
3. Pulse el botón **Aceptar**.

En la Figura 4.3 puede observar gráficamente esta secuencia de pasos.



Es recomendable que las variables que se generen a través de la calculadora tengan el mismo nombre de la operación, ya que esto facilita la elaboración de gráficos para el modelo de regresión.

Figura 4.2
ERLA 2011 Estadística ICM/ESPOL
"Regresión lineal simple"

The figure illustrates the process of performing a simple linear regression in the ERLA 2011 software. It shows the menu navigation and the configuration of the regression model, followed by the resulting statistical report.

Regresión lineal Dialog Box:

- Variables:** CalIESPOL, P1, P2, P3, P4, P5, P6, P7, P8, P11, P12, P13, P14, P15, P16, P17, P18, P19
- Variable a ser explicada:** P10
- Variable de explicación:** P9

Reporte 4 - Regresión Lineal:

Regresión Lineal
 $P10 = 1.31 + 0.8 * P9$

R2: 0.61 | R2 Ajustado: 0.61 | Desviación Estándar: 1.2

Análisis de Varianza:

Tabla ANOVA					
	G.L.	S.C.	M.C.	F	P
REGRESIÓN	1	830.57	830.57	575.21	0
ERROR	371	535.7	1.44		
TOTAL	372	1366.27			

Tabla de Inferencia respecto a los parámetros betas

Tabla de Estimadores				
	ESTIMADOR	E.E.ESTIMADOR	T	P
CONSTANTE	1.31	0.28	4.74	0
P9	0.8	0.03	23.98	0

Intervalos de Confianza para los parámetros betas

95% de Confianza		
	INFERIOR	SUPERIOR
CONSTANTE	0.77	1.86
P9	0.74	0.87

Interpretación de resultados

Cada vez que “ejecute” un modelo de Regresión Lineal, usted obtendrá:

- Modelo, el cual para el ejemplo es “ $P_{10}=1.31+0.8 \cdot P_9$ ” donde P_9 y P_{10} son las proposiciones (*Identifico a los estudiantes de la ESPOL por su responsabilidad*) y (*Identifico a los estudiantes de la ESPOL por su honestidad*) del estudio “Imagen de la ESPOL es Guayaquil” las cuales fueron calificadas en una escala del 1 al 10, donde 1 indica “Total desacuerdo” con lo planteado, y 10 “Total acuerdo”. Es decir que:

$$Y_i = b_0 + b_1 X_i = P_{10}=1.31+0.8 \cdot P_9$$

- Potencia de explicación del modelo (R^2 y R^2 Ajustado)
- Desviación estándar (s)
- Tabla ANOVA con todos sus valores característicos: Fuentes de variación (1era columna), grados de libertad ($G.L.$), sumas y medias cuadráticas ($S.C.$ y $M.C.$, respectivamente), el estadístico de prueba F (F) y el valor p (P)..
- Tabla de inferencia respecto a los parámetros betas. El valor del estimador (*ESTIMADOR*), el error estándar del estimador (*E. E. ESTIMADOR*), el estadístico de prueba t (T) y el valor p (P)
- Intervalos de confianza para los parámetros betas utilizando un nivel de confianza del 95%. Se puede distinguir en los resultados de la tabla el límite inferior (*INFERIOR*) y el límite superior (*SUPERIOR*).

Figura 4.3
ERLA 2011 Estadística ICM/ESPOL
"Regresión lineal simple polinómica"

The screenshot shows the ERLA 2011 software interface. The 'Regresión' menu is open, and the 'Regresión lineal' option is selected. The 'Regresión lineal' dialog box is displayed, showing the following configuration:

- Variables:** A list of variables including CalIESPOL, P1, P2, P3, P4, P5, P6, P7, P8, P11, P12, P13, P14, P15, P16, P17, P18, and P19. A green circle with the number '2' is placed over this list.
- Variable a ser explicada:** P10
- Variable de explicación:** P9, P9.^2

Below the dialog box, the 'Reporte 1' window displays the results of the regression analysis:

Regresión Lineal

$$P10 = 1.12 + 0.86 * P9 + 0 * P9.^2$$

R2: 0.61 | R2 Ajustado: 0.61 | Desviación Estándar: 1.2

Análisis de Varianza:

Tabla ANOVA					
	G.L.	S.C.	M.C.	F	P
REGRESIÓN	2	830.8	415.4	287.04	0
ERROR	370	535.46	1.45		
TOTAL	372	1366.27			

Tabla de Inferencia respecto a los parámetros betas

Tabla de Estimadores				
	ESTIMADOR	E.E.ESTIMADOR	T	P
CONSTANTE	1.12	0.54	2.07	0.04
P9	0.86	0.16	5.46	0
P9.^2	-0	0.01	-0.41	0.69

Intervalos de Confianza para los parámetros betas

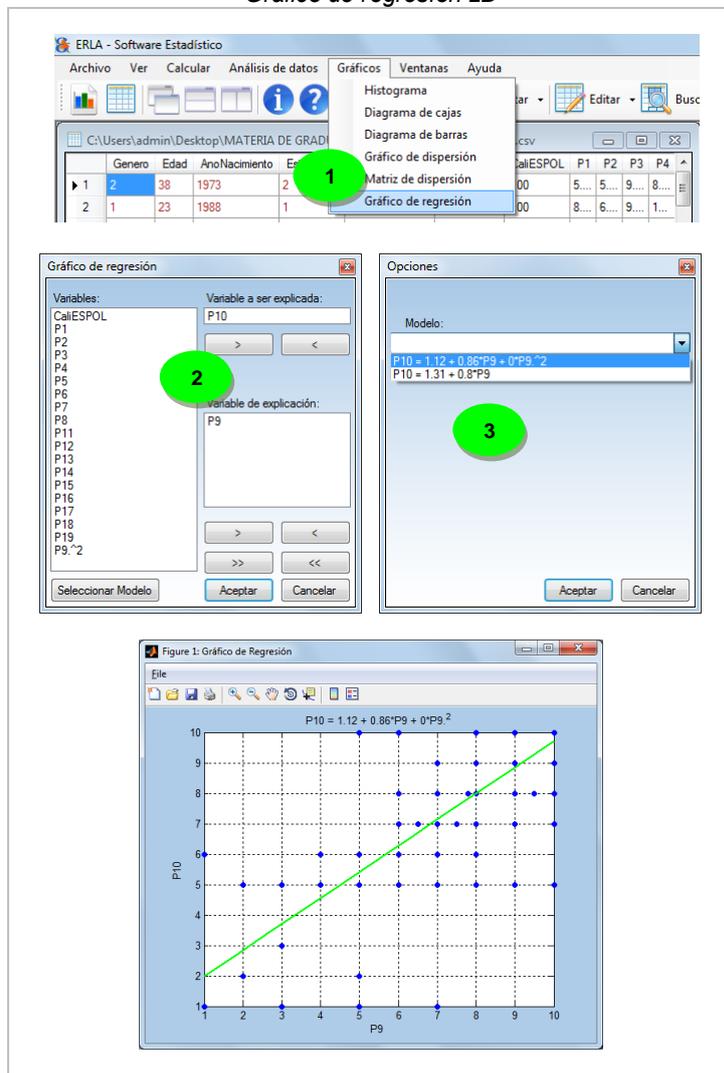
95% de Confianza		
	INFERIOR	SUPERIOR
CONSTANTE	0.06	2.19
P9	0.55	1.17
P9.^2	-0.03	0.02

Gráfico de regresión

Una vez que se haya obtenido el modelo, es posible visualizar el gráfico correspondiente. Para ello vaya a:

1. Barra de menús ➤ Gráficos ➤ Gráfico de regresión
2. Seleccione las variables del modelo que desee graficar, en este caso P_{10} como variable a ser explicada y P_9 como variable de explicación.
3. Pulse el botón **Seleccionar modelo**.
4. Seleccione el modelo que desee de la lista y pulse el botón **Aceptar**. Véase Figura 4.4.

Figura 4.4
ERLA 2011 Estadística ICM/ESPOL
"Gráfico de regresión 2D"



Regresión lineal múltiple

En la matriz de correlación se pueden observar que la proposición *P12* (La capacitación que se imparte en la ESPOL permite un adecuado desempeño a nivel Profesional) está altamente correlacionada con *P13* (Ser graduado en la ESPOL da ventajas competitivas frente a sus potenciales empleadores) y *P14* (Ser graduado en la ESPOL concede ventajas frente a sus competidores de otras universidades). Veamos que sucede si se intenta explicar *P12* en función de *P13* y *P14*. Los pasos son:

1. Barra de menús > **Análisis de datos** > **Regresión** > **Regresión lineal**
2. Seleccione la variable a ser explicada, y las variables de explicación, para el ejemplo *P12* como variable a ser explicada, *P13* y *P14* como variables de explicación.
3. Pulse el botón **Aceptar**.

En la Figura 4.5 se pueden observar los resultados.

Regresión lineal múltiple polinómica con interacciones

De manera similar al caso de la regresión lineal simple polinómica, es necesario usar la **Calculadora** de ERLA para generar las nuevas variables. Para el ejemplo veamos qué pasa con el modelo de regresión lineal múltiple si le agregamos términos cuadráticos y la interacción entre variables. Siga los siguientes pasos:

1. Barra de menús > **Calcular** > **Calculadora**
2. Realice la operación necesaria, en este caso $P13.^2$ y deje el cuadro de texto **Nombre de la variable** en blanco para que tome el nombre de la operación.
3. Pulse el botón **Aceptar**.
4. Repita el procedimiento para generar las variables $P14.^2$ y $P13.*P14$.

Las operaciones de la **Calculadora** deben expresarse en el lenguaje de MATLAB. Véase Figura 4.6.

Figura 4.5
ERLA 2011 Estadística ICM/ESPOL
"Regresión lineal múltiple"

The figure illustrates the process of performing multiple regression in ERLA 2011. It is divided into three main sections:

1. Software Interface: The top screenshot shows the 'ERLA - Software Estadístico' window. The 'Análisis de datos' menu is open, and 'Regresión' is selected. The 'Regresión lineal' option is highlighted in the sub-menu. A green circle with the number '1' is placed over the 'Regresión' menu item.

2. Regression Dialog Box: The middle screenshot shows the 'Regresión lineal' dialog box. In the 'Variables:' list, 'CalIESPOL' is selected. In the 'Variable a ser explicada:' field, 'P12' is entered. In the 'Variable de explicación:' list, 'P13' and 'P14' are selected. A green circle with the number '2' is placed over the 'CalIESPOL' variable in the list.

3. Regression Report: The bottom screenshot shows the 'Reporte 1' window. It displays the following information:

Regresión Lineal

$$P12 = 1.9 + 0.4 \cdot P13 + 0.38 \cdot P14$$

R2: 0.6 | R2 Ajustado: 0.6 | Desviación Estándar: 1.01

Análisis de Varianza:

Tabla ANOVA					
	G.L.	S.C.	M.C.	F	P
REGRESIÓN	2	576.89	288.44	281.08	0
ERROR	370	379.7	1.03		
TOTAL	372	956.59			

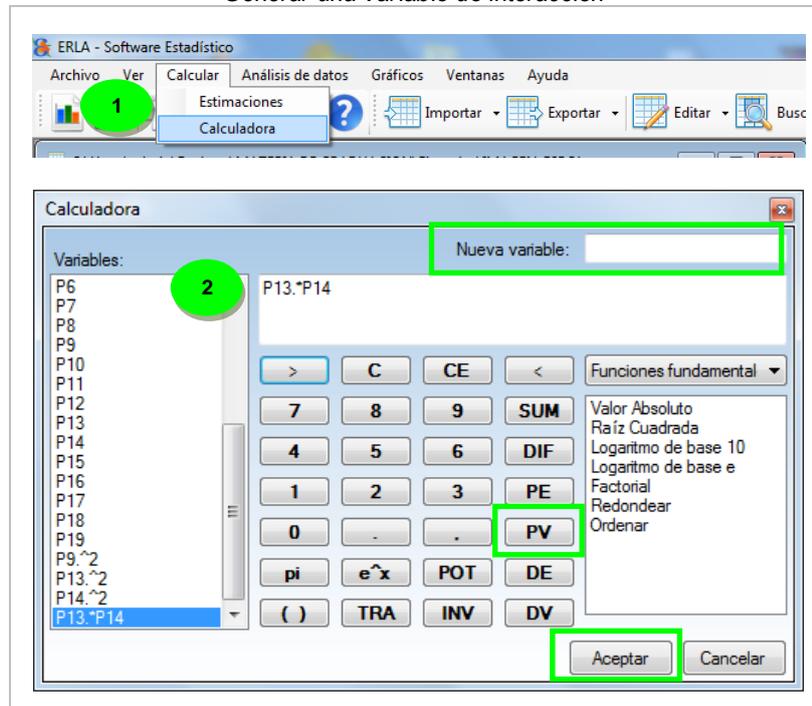
Tabla de Inferencia respecto a los parámetros betas

Tabla de Estimadores				
	ESTIMADOR	E.E.ESTIMADOR	T	P
CONSTANTE	1.9	0.28	6.7	0
P13	0.4	0.05	8.28	0
P14	0.38	0.05	8.07	0

Intervalos de Confianza para los parámetros betas

95% de Confianza		
	INFERIOR	SUPERIOR
CONSTANTE	1.35	2.46
P13	0.3	0.49
P14	0.29	0.48

Figura 4.6
ERLA 2011 Estadística ICM/ESPOL
“Generar una variable de interacción”



Para obtener el modelo de regresión lineal múltiple polinómica con Interacciones en ERLA, vaya a:

1. **Barra de menus** > **Análisis de datos** > **Regresión** > **Regresión lineal**
2. Seleccione la variable a ser explicada. y las variables de explicación, que para el ejemplo son $P13$, $P14$, $P13.^2$, $P14.^2$ y $P13.*P14$
3. Pulse el botón **Aceptar**.

Nótese que la potencia de explicación del modelo aumentó de un 60% a 64%.

Figura 4.7
 ERLA 2011 Estadística ICM/ESPOL
 “Regresión lineal múltiple polinómica con interacciones”

The figure illustrates the steps to perform a multiple polynomial regression with interactions in ERLA 2011. It shows the software interface, the selection of variables in the regression dialog, and the resulting statistical report.

Regresión lineal Dialog Box:

- Variables:** CaliESPOL, P1, P2, P3, P4, P5, P6, P7, P8, P9, P10, P11, P15, P16, P17, P18, P19, P9.^2
- Variable a ser explicada:** P12
- Variable de explicación:** P13, P14, P13.^2, P14.^2, P13.*P14

Reporte 1 - Regresión Lineal:

$$P12 = 0.49 + 0.33 * P13 + 0.78 * P14 + 0.07 * P13.^2 + 0.03 * P14.^2 - 0.12 * P13.*P14$$

R2: 0.64 | R2 Ajustado: 0.64 | Desviación Estándar: 0.98

Análisis de Varianza:

Tabla ANOVA					
	G.L.	S.C.	M.C.	F	P
REGRESIÓN	5	616.94	123.39	133.32	0
ERROR	367	339.65	0.93		
TOTAL	372	956.59			

Tabla de Inferencia respecto a los parámetros betas:

Tabla de Estimadores					
	ESTIMADOR	E.E.ESTIMADOR	T	P	
CONSTANTE	0.49	0.61	0.79	0.43	
P13	0.33	0.22	1.49	0.14	
P14	0.78	0.22	3.64	0	
P13.^2	0.07	0.02	4.03	0	
P14.^2	0.03	0.02	1.72	0.09	
P13.*P14	-0.12	0.02	-8.27	0	

Intervalos de Confianza para los parámetros betas:

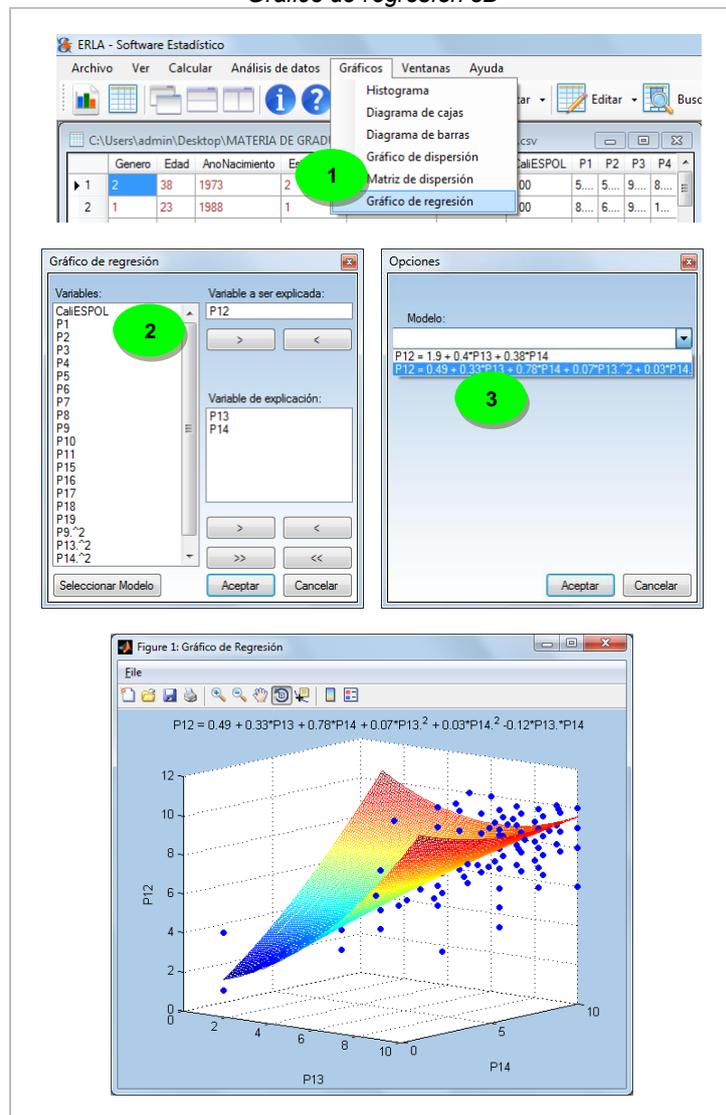
95% de Confianza		
	INFERIOR	SUPERIOR
CONSTANTE	-0.72	1.7
P13	-0.11	0.77
P14	0.36	1.21
P13.^2	0.03	0.1
P14.^2	-0	0.07
P13.*P14	-0.15	-0.08

Gráfico de regresión

Para graficar el modelo obtenido realice lo siguiente:

1. Barra de menús > Gráficos > Gráfico de regresión
2. Seleccione las variables del modelo que desee graficar, en este caso $P12$ como variable a ser explicada y las variables $P13$ y $P14$ como variables de explicación.
3. Pulse el botón **Seleccionar modelo**.
4. Seleccione el modelo que desee de la lista desplegable "Modelo" y pulse el botón **Aceptar**. Véase Figura 4.8.

Figura 4.8
ERLA 2011 Estadística ICM/ESPOL
"Gráfico de regresión 3D"





En el cuadro de diálogo **Gráfico de regresión** se permite un máximo de dos variables de explicación, ya que un modelo que involucre tres o más variables de explicación no puede ser graficado.

Los términos polinómicos y las interacciones entre variables no son consideradas variables diferentes en el gráfico de regresión; siempre y cuando lleven por nombre la expresión que permite generarlas.

Estimaciones

Para realizar estimaciones de cualquier modelo que haya sido generado en ERLA, vaya a:

1. **Barra de menues > Calcular > Estimaciones**
2. Seleccione un modelo de la lista **Modelo**.
3. Ingrese los valores de las variables de explicación y luego pulse el botón **Aceptar**.

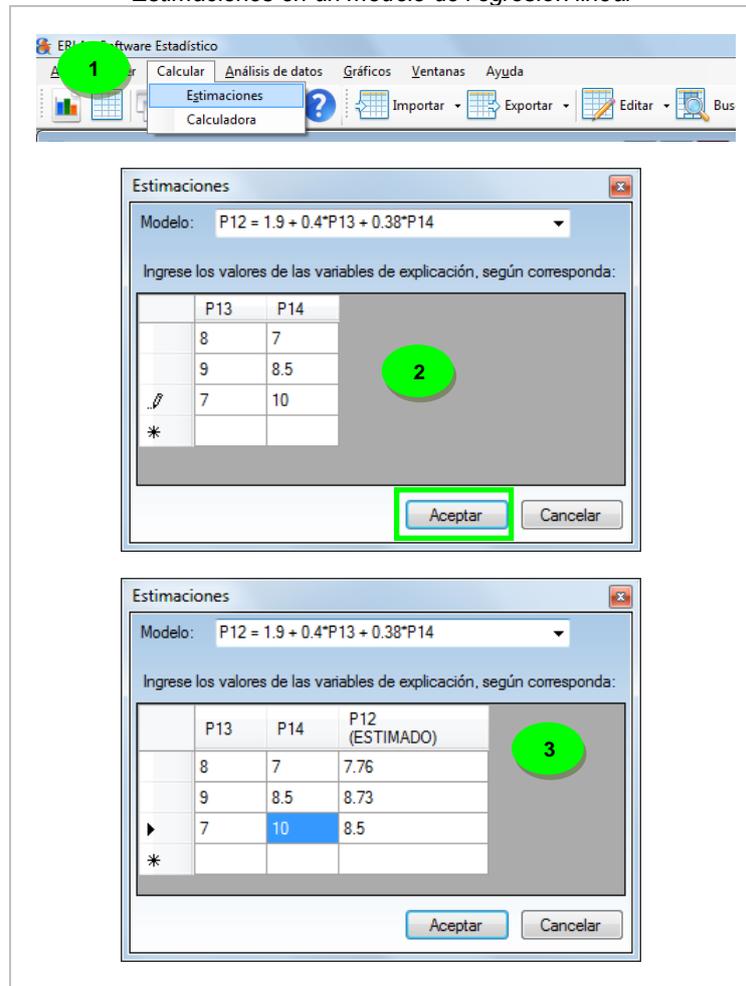
Los resultados aparecerán en una columna a la derecha tal como se muestra en la Figura 4.9.



Cada modelo de regresión lineal que sea generado en ERLA aparecerá en la lista **Modelo** del cuadro de diálogo **Estimaciones**.

Una vez que se cierra el programa, estos modelos son eliminados de la lista.

Figura 4.9
ERLA 2011 Estadística ICM/ESPOL
"Estimaciones en un modelo de regresión lineal"



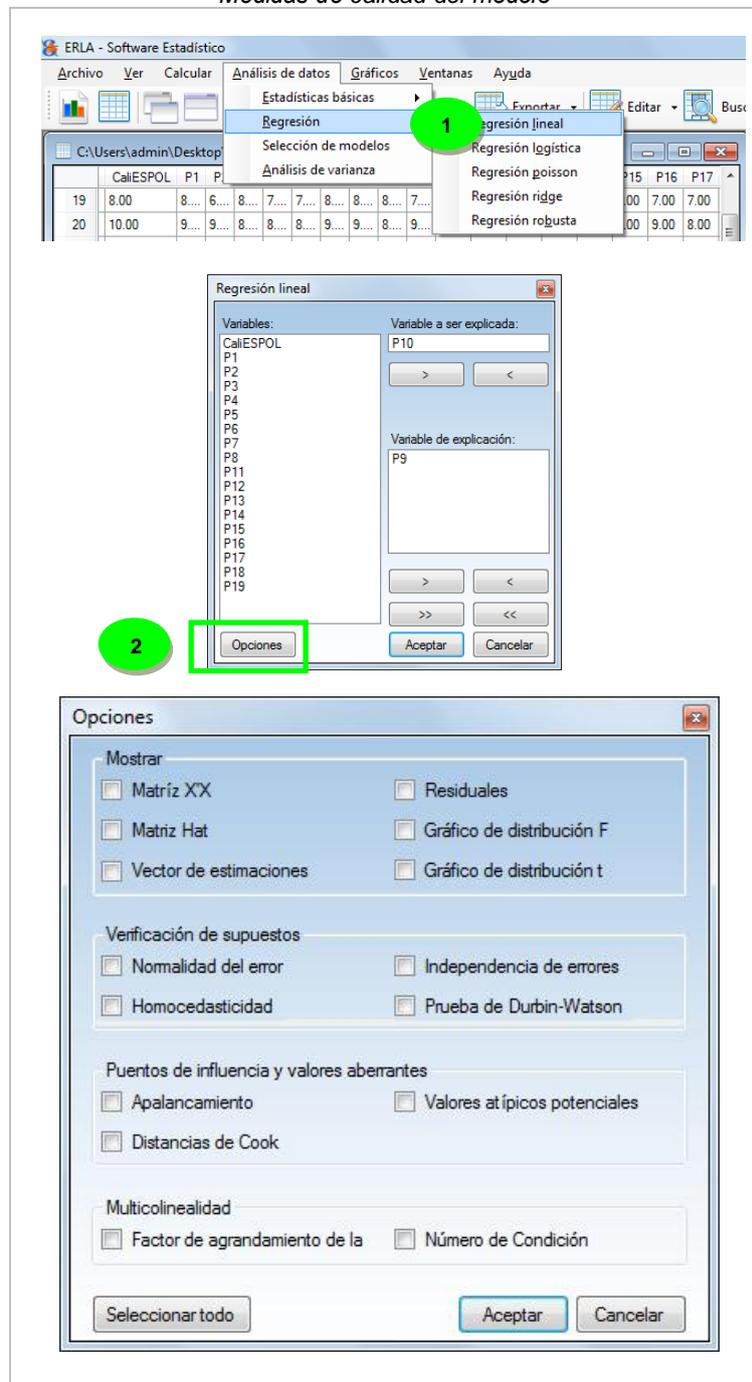
Verificación de la calidad en modelos de regresión lineal

Existen distintos métodos que permiten determinar la bondad de un modelo de regresión, ERLA incluye tales opciones en el cuadro de diálogo **Regresión lineal** al cual se puede acceder mediante la secuencia:

1. **Barra de menú** > **Análisis de datos** > **Regresión** > **Regresión lineal**

2. Pulse el botón **Opciones** del cuadro de diálogo **Regresión lineal**. Véase Figura 4.10.

Figura 4.10
ERLA 2011 Estadística ICM/ESPOL
"Medidas de calidad del modelo"



Es decir, una vez seleccionado el modelo de regresión lineal, debe pulsar el botón **Opciones** para visualizar los estadísticos que ofrece ERLA para verificar la validez del modelo. Las opciones relacionadas con la verificación de la calidad del modelo están principalmente en los grupos:

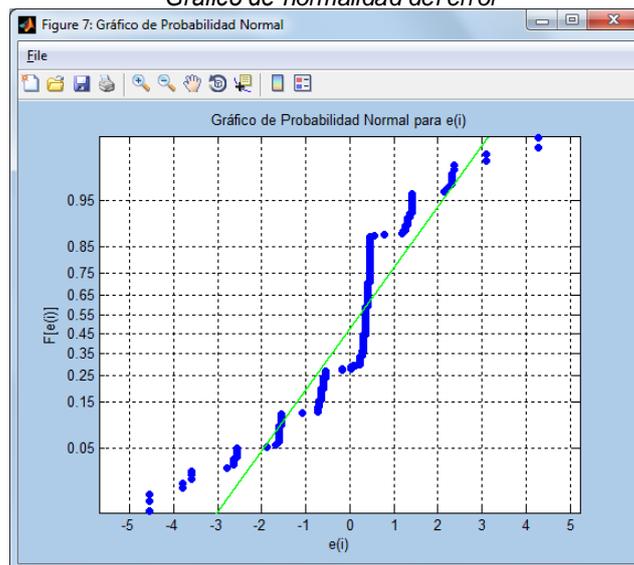
1. Verificación de supuestos.
2. Puntos de Influencia y valores aberrantes.
3. Multicolinealidad.

Verificación de supuestos del error

Gráfico de normalidad del error

ERLA ofrece el gráfico de probabilidad normal como una opción para verificar el cumplimiento del supuesto de normalidad del error. Marque la opción **Normalidad del error** para generar este gráfico. Véase Figura 4.11.

Figura 4.11
ERLA 2011 Estadística ICM/ESPOL
"Gráfico de normalidad del error"



*Gráfico de
homocedasticidad*

Para verificar si la varianza del error es constante, ERLA ofrece el gráfico de residuales versus valores ajustados o estimados de la variable a ser explicada. Seleccione la opción **Homocedasticidad** para obtener este gráfico. Véase Figura 4.12.

Figura 4.12
ERLA 2011 Estadística ICM/ESPOL
"Gráfico de homocedasticidad"

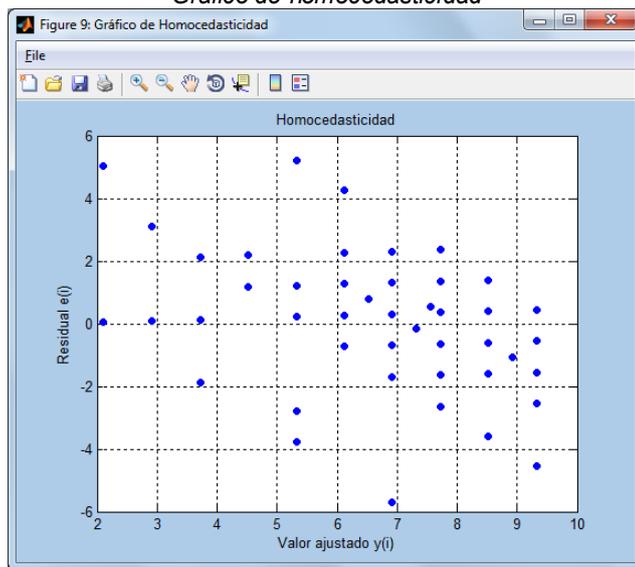
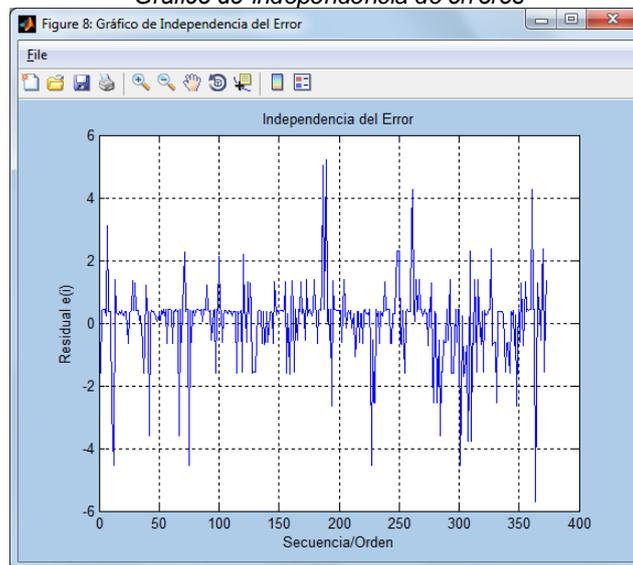


Gráfico de independencia de errores

Para verificar si los errores no están correlacionados, ERLA ofrece el gráfico de los residuales versus la secuencia/orden. Active la opción **Independencia de errores**. Véase Figura 4.13.

Figura 4.13
ERLA 2011 Estadística ICM/ESPOL
"Gráfico de independencia de errores"



Prueba Durbin-Watson

Como resultados de esta prueba obtendrá el valor del estadístico de Durbin-Watson con su respectivo valor p para una prueba de dos colas. Para realizar esta prueba seleccione la opción **Prueba de Durbin-Watson**. Véase Figura 4.14.

Figura 4.14
ERLA 2011 Estadística ICM/ESPOL
"Prueba Durbin-Watson"

Prueba de Durbin-Watson

Estadístico D-W: 1.75

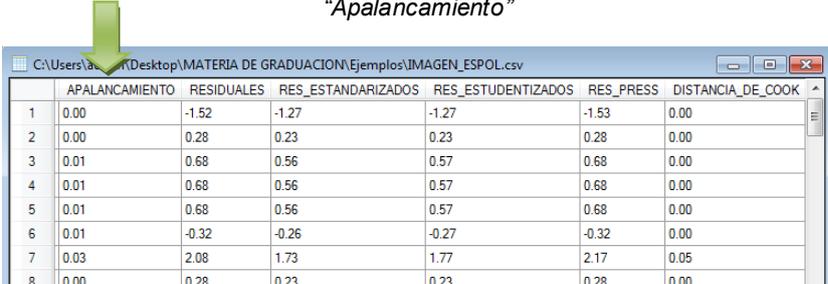
Valor p : 0.01

Puntos de Influencia y Valores Aberrantes

Apalancamiento

El vector de apalancamientos se presenta como una nueva variable en la **ventana de datos**, junto a las variables consideradas en el modelo de regresión lineal. Active la opción **Apalancamiento**. Véase Figura 4.15.

Figura 4.15
ERLA 2011 Estadística ICM/ESPOL
"Apalancamiento"

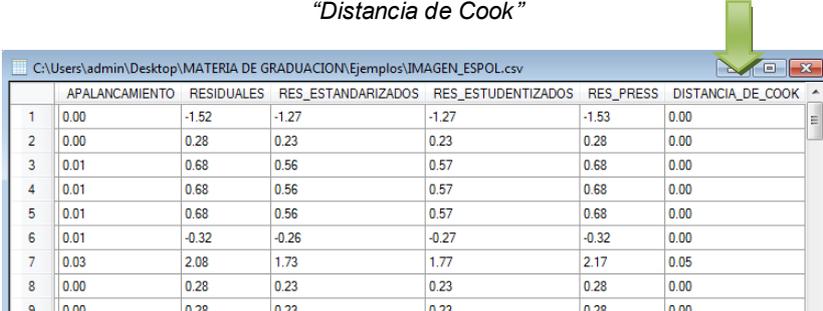


	APALANCAMIENTO	RESIDUALES	RES_ESTANDARIZADOS	RES_ESTUDENTIZADOS	RES_PRESS	DISTANCIA_DE_COOK
1	0.00	-1.52	-1.27	-1.27	-1.53	0.00
2	0.00	0.28	0.23	0.23	0.28	0.00
3	0.01	0.68	0.56	0.57	0.68	0.00
4	0.01	0.68	0.56	0.57	0.68	0.00
5	0.01	0.68	0.56	0.57	0.68	0.00
6	0.01	-0.32	-0.26	-0.27	-0.32	0.00
7	0.03	2.08	1.73	1.77	2.17	0.05
8	0.00	0.28	0.23	0.23	0.28	0.00

Distancia de Cook

El vector que contiene las distancias de Cook aparece como una nueva variable en la **ventana de datos** junto a las variables consideradas en el modelo de regresión lineal. Habilite la opción **Distancia de Cook**. Véase Figura 4.16.

Figura 4.16
ERLA 2011 Estadística ICM/ESPOL
"Distancia de Cook"



	APALANCAMIENTO	RESIDUALES	RES_ESTANDARIZADOS	RES_ESTUDENTIZADOS	RES_PRESS	DISTANCIA_DE_COOK
1	0.00	-1.52	-1.27	-1.27	-1.53	0.00
2	0.00	0.28	0.23	0.23	0.28	0.00
3	0.01	0.68	0.56	0.57	0.68	0.00
4	0.01	0.68	0.56	0.57	0.68	0.00
5	0.01	0.68	0.56	0.57	0.68	0.00
6	0.01	-0.32	-0.26	-0.27	-0.32	0.00
7	0.03	2.08	1.73	1.77	2.17	0.05
8	0.00	0.28	0.23	0.23	0.28	0.00
9	0.00	0.28	0.23	0.23	0.28	0.00

Valores atípicos potenciales

En la **ventana de reportes** de ERLA se presentan los valores atípicos potenciales en caso de que existan con sus correspondientes residuales, caso contrario, se presentará un mensaje diciendo que en el modelo no hay valores atípicos potenciales. Para mostrar estos valores, seleccione la opción **Valores atípicos potenciales**. Véase Figura 4.17.

Figura 4.17
ERLA 2011 Estadística ICM/ESPOL
"Valores atípicos potenciales"

Valores atípicos potenciales

Las siguientes observaciones son consideradas valores atípicos potenciales:

Valores atípicos potenciales				
Observaciones	Residuales	Residuales Estandarizados	Residuales Estudentizados	Residuales Press
12	-4.32	-3.6	-3.67	-4.5
75	-4.32	-3.6	-3.67	-4.5
187	3.89	3.23	3.35	4.18
189	4.68	3.9	4	4.93
227	-4.32	-3.6	-3.67	-4.5
261	3.88	3.23	3.29	4.02
301	-4.32	-3.6	-3.67	-4.5
307	-4.32	-3.59	-3.68	-4.52
310	-4.32	-3.59	-3.68	-4.52
361	3.88	3.23	3.29	4.02
364	-5.92	-4.93	-5.11	-6.36

ERLA considera una observación como atípica potencial si el correspondiente residual (cualquiera de los antes mencionados), es mayor a 3 desviaciones estándar de su respectivo residual.

Multicolinealidad

Factor de inflación o agrandamiento de la varianza

En la **ventana de reporte** se presenta el R^2 y el FIV (Factor de inflación de la varianza) para cada una de las variables de explicación incluidas en el modelo. Active la opción **Factor de agrandamiento de la varianza**. Véase Figura 4.18.

Figura 4.18
ERLA 2011 Estadística ICM/ESPOL
"Factor de agrandamiento de la varianza"

<i>Factor de Agrandamiento de la Varianza</i>		
	R2	FAV
P13	0.58	2.39
P14	0.58	2.39

Número de condición

El valor correspondiente al número de condición se presenta en la **ventana de reporte**. Para esto, habilite la opción **Número de condición**. Véase Figura 4.19.

Figura 4.19
ERLA 2011 Estadística ICM/ESPOL
"Número de condición"

<i>Número de Condición</i>
Número de condición = 66.38

Selección de modelos

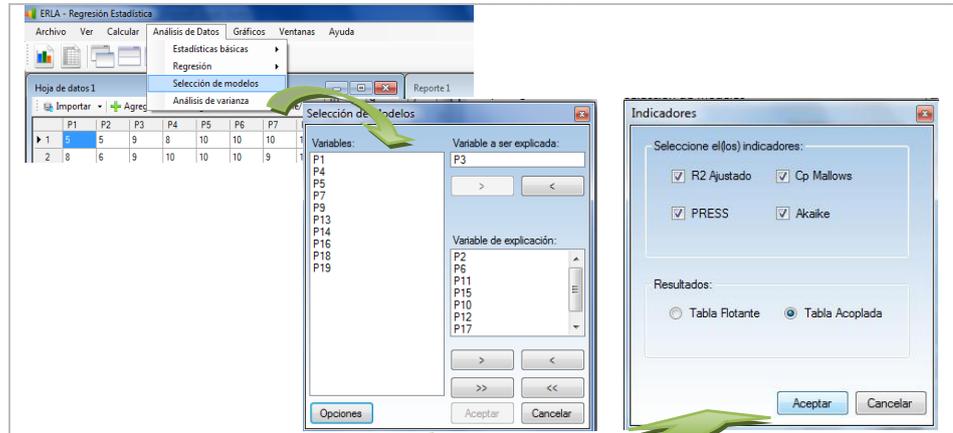
El ítem selección de modelos, permite al usuario determinar el mejor modelo de regresión, basado en los indicadores R^2 Ajustado, C_p Mallows, Akaike y el estadístico PRESS.

De acuerdo con el ejemplo propuesto, “Imagen de la ESPOL en Guayaquil”, para este caso se considera el conjunto de proposiciones de las cuales como variable a ser explicada se escoge la proposición 3 y como variables de explicación las proposiciones 2, 6, 11, 15, 10, 12, 17 y 18. Véase ANEXO I.

Los pasos a seguir en ERLA son:

1. **Barra de menú** > **Análisis de datos** > **Selección de modelos**
2. En el cuadro de diálogo **Selección de modelos**, seleccione la variable a ser explicada y las variables de explicación y pulse el botón **Opciones**.
3. En el cuadro de diálogo **Indicadores** seleccione el estadístico mediante el cual desea hacer la comparación y pulse el botón **Aceptar**. Los resultados se mostraran en la **ventana de reporte**. Véase Figura 4.20.

Figura 4.20
ERLA 2011 Estadística ICM/ESPOL
"Selección de Modelos"



Resultado Tabla Acoplada

Reporte 1

Reporte generado el: 3/21/2011 5:00:56 AM

Modelos de Selección

Modelo R2 Ajustado

R2 Ajustado:	
	R2 Ajustado
/P2	0.0736
/P6	0.2924
/P11	0.7068
/P15	0.3376
/P10	0.4097
/P12	0.5519
/P17	0.1852
/P8	0.3192
/P2/P6	0.2906
/P2/P11	0.7088
/P2/P15	0.3421
/P2/P10	0.4176

Resultado de Tabla Flotante

File	R2 Ajustado
/P2	0.0736
/P6	0.2924
/P11	0.7068
/P15	0.3376
/P10	0.4097
/P12	0.5519
/P17	0.1852

File	Cp Mallows
/P2	578.4735
/P6	395.9738
/P11	50.4500
/P15	358.3619
/P10	298.1889
/P12	179.6727
/P17	466.2977

File	AKAIKE
/P2	250.2364
/P6	196.8715
/P11	22.4202
/P15	183.8283
/P10	160.9908
/P12	106.4454
/P17	724.8105

File	PRESS
/P2	2.8160e+03
/P6	821.0808
/P11	240.8142
/P15	869.1404
/P10	668.7162
/P12	391.9255
/P17	1.4708e+03

Los resultados en la “*Ventana de reporte*” se presenta de dos formas:

Tabla acoplada:

- La matriz de resultados, donde la primer columna representa al número total de combinaciones posibles que se generan a partir del numero de variables que seleccione el usuario en el cuadro de dialogo y las columnas restantes serán igual al número de indicadores que opte el usuario con los valores respectivos de cada indicador.

Tabla Flotante:

- Aparecen los valores de cada indicador en ventanas separadas.
- La combinación óptima de variables para el caso del coeficiente de determinación será el de mayor valor.
- La combinación óptima de variables para el caso del indicador Akaike es el de menor valor.
- La combinación óptima de variables para el caso del indicador Cp Mallows será el valor menor o igual al número de parámetros en dicha combinación.

La combinación óptima de variables para el caso del indicador PRESS es el de menor valor.

Regresión logística

La regresión logística es útil en los casos en los que se desee predecir la presencia o ausencia de una determinada característica según un conjunto de variables predictivas. Así la ausencia o presencia de esa característica se denota mediante el valor de uno (presencia) o el cero (ausencia).

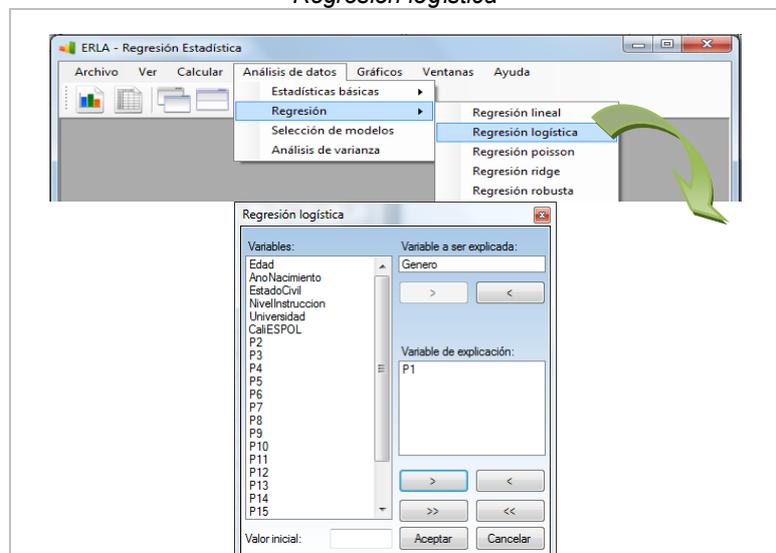
El modelo de regresión logística de probabilidades $\pi(x_i)$ es:

$$\pi(x_i) = \frac{e^{x_i' \beta}}{1 + e^{x_i' \beta}}; \text{ Donde } x_i' \beta = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}$$

Una vez que se hayan ingresado los datos, seguimos una serie de pasos para realizar la regresión logística:

1. Barra de menú **Análisis de datos** **Regresión** **Regresión logística**
2. En el cuadro de diálogo **Regresión logística** seleccione las variables a utilizar.
3. Pulse el botón **Aceptar**. Véase Figura 4.21.

Figura 4.21
ERLA 2011 Estadística ICM/ESPOL
"Regresión logística"



Para el ejemplo, la variable a ser explicada será la pregunta 5 del formulario y la variable de explicación, la proposición 1. Una vez seleccionadas las variables, se pulsa el botón **Aceptar**. Los resultados se presentan en la **ventana de reporte** junto con el gráfico correspondiente. Véase Figura 4.22.

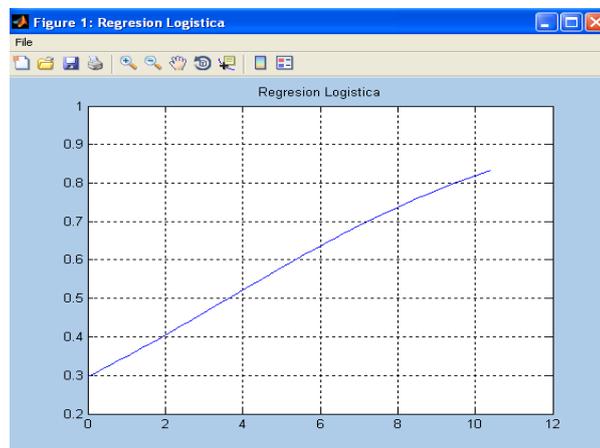
Figura 4.22

ERLA 2011 Estadística ICM/ESPOL
"Regresión logística: Resultados"

Salida en
ventana de
reporte



Salida en
ventana de
gráfico



**Interpretación
de resultados**

En la ventana de reportes se presenta dos tablas la primera para estimar los parámetros betas y el cálculo de intervalos de confianza para cada parámetro, por lo tanto a continuación se interpreta los siguientes resultados:

Dado el modelo de regresión logística

$$\pi(x_i) = \frac{e^{-0.8592+0.2628x_1}}{1+e^{-0.8592+0.2628x_1}},$$

se obtuvo como resultado lo que se muestra en la tabla de estimación de parámetros donde se obtiene que el estimador del parámetro β_0 es -0.8292 y de β_1 es 0.2367. Además se presenta el Error Estándar para β_0 es 0.2628 y β_1 0.0422 los cuales permitirán el cálculo del estadístico de prueba T que también se presenta en la Tabla con sus respectivos estimadores de los parámetros de beta donde β_0 es -3.27 y β_1 es 5.608 y con estos resultados rechazar o no rechazar la hipótesis planteada de acuerdo al valor p que se obtuvo lo cual para este caso el nivel de confianza fue de 0.012 donde decimos que con el $(1-\alpha)\%$ de confianza rechazamos la hipótesis nula a favor de la hipótesis alterna ya que el valor p es 0.012 lo cual indica que el modelo planteado no es el correcto.

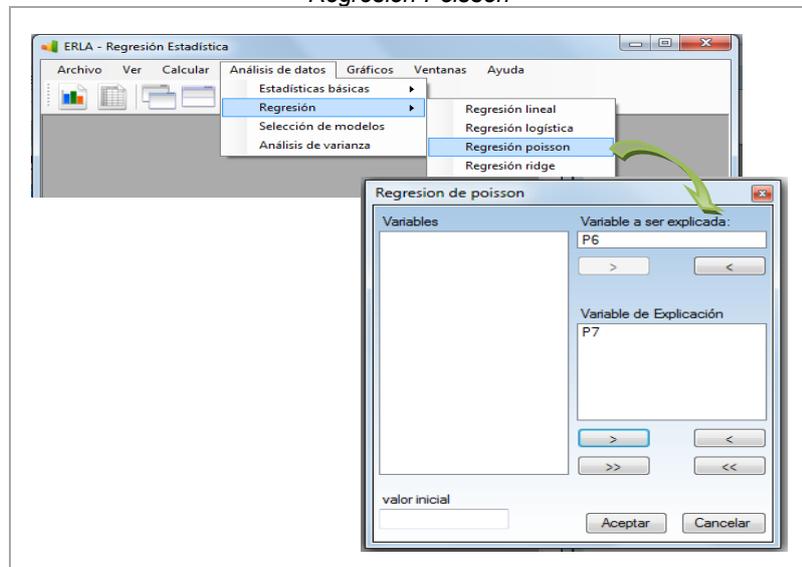
En la Tabla de Intervalos de Confianza, se muestran los intervalos de confianza para ambos estimadores del parámetro beta, así está el límite inferior y superior de cada uno con un nivel de confianza del 95%, los cuales dieron con resultado para donde β_0 -1.376 y β_1 0.1537 en los límites inferior y β_0 -0.3424 y β_1 0.3198 en los límites superior.

Regresión Poisson

Para hacer la regresión Poisson se realizan los siguientes pasos:

1. **Barra de menú** > **Análisis de datos** > **Regresión** > **Regresión Poisson**
2. En el cuadro de diálogo **Regresión Poisson** seleccione las variables a utilizar.
3. Pulse el botón **Aceptar**. Véase Figura 4.23.

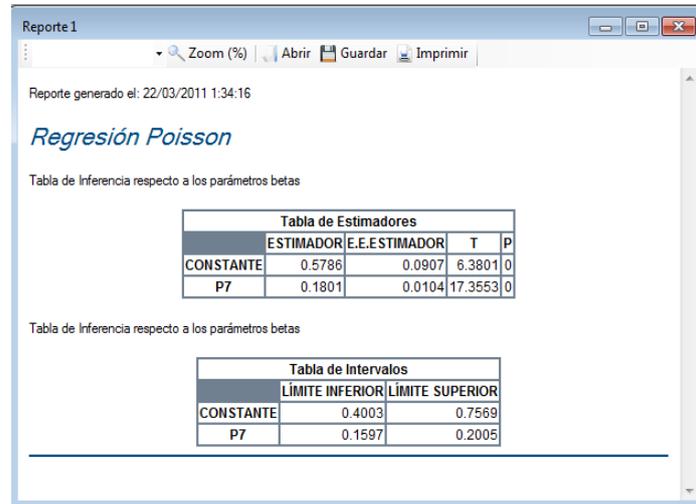
Figura 4.23
ERLA 2011 Estadística ICM/ESPOL
"Regresión Poisson"



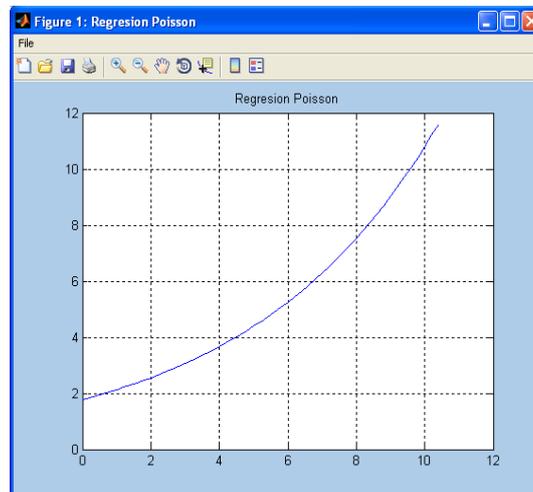
Para el ejemplo, la variable a ser explicada será la proposición 6 y la variable de explicación la proposición 7. Una vez seleccionadas las variables, se pulsa el botón **Aceptar**. Los resultados se presentan en la **ventana de reporte** junto con el gráfico correspondiente. Véase Figura 4.25.

Salida de
ventana de
reporte

Figura 4.24
ERLA 2011 Estadística ICM/ESPOL
"Regresión Poisson: Resultados"



Salida de
Ventana de
Gráfico



Interpretación de resultados

Para la regresión de Poisson se realizó las mismas tablas de regresión logística para mostrar los resultados donde se diferencia el modelo a plantearse, el cual es el siguiente:

$Y = e^{(0.5786+0.1801x)}$, se obtuvo como resultado lo que se muestra en la tabla de estimación de parámetros donde se obtiene que el estimador del parámetro β_0 es 0.5786 y de β_1 es 0.1801. Además se presenta el Error Estándar para β_0 es 0.0907 y β_1 0.0104, los cuales permitirán el cálculo del estadístico de prueba T que también se presenta en la Tabla con sus respectivos estimadores de los parámetros de beta donde β_0 es 6.3801 y β_1 es 17.3553 y con estos resultados rechazar o no rechazar la hipótesis planteada de acuerdo al valor p que se obtuvo lo cual para este caso el nivel de confianza fue de 0 donde decimos que con el $(1-\alpha)\%$ de confianza rechazamos la hipótesis nula a favor de la hipótesis alterna ya que el valor p es 0 lo cual indica que el modelo planteado no es el correcto.

En la Tabla de Intervalos de Confianza, se muestran los intervalos de confianza para ambos estimadores del parámetro beta, así está el límite inferior y superior de cada uno con un nivel de confianza del 95%, los cuales dieron con resultado para donde β_0 0.40003 y β_1 0.1597 en los límites inferior y β_0 0.7569 y β_1 0.2005 en los límites superior.

Regresión Ridge

Para ilustrar el uso de regresión Ridge tomaremos un ejemplo del libro “*Probabilidad y Estadística Fundamentos y Aplicaciones*” de G. Zurita el cual se ajusta a las necesidades de nuestro problema:

Los integrantes de una cooperativa arrocera están interesados en que se les ayude a conocer el efecto de la humedad, en cm/m^2 y la temperatura, en grados Celsius, sobre la producción de arroz (variable dependiente), en toneladas métricas por hectárea. Los datos se presentan en la Tabla 4.1.

Tabla 4.1
ERLA 2011 Estadística ICM/ESPOL
“Datos para regresión Ridge”

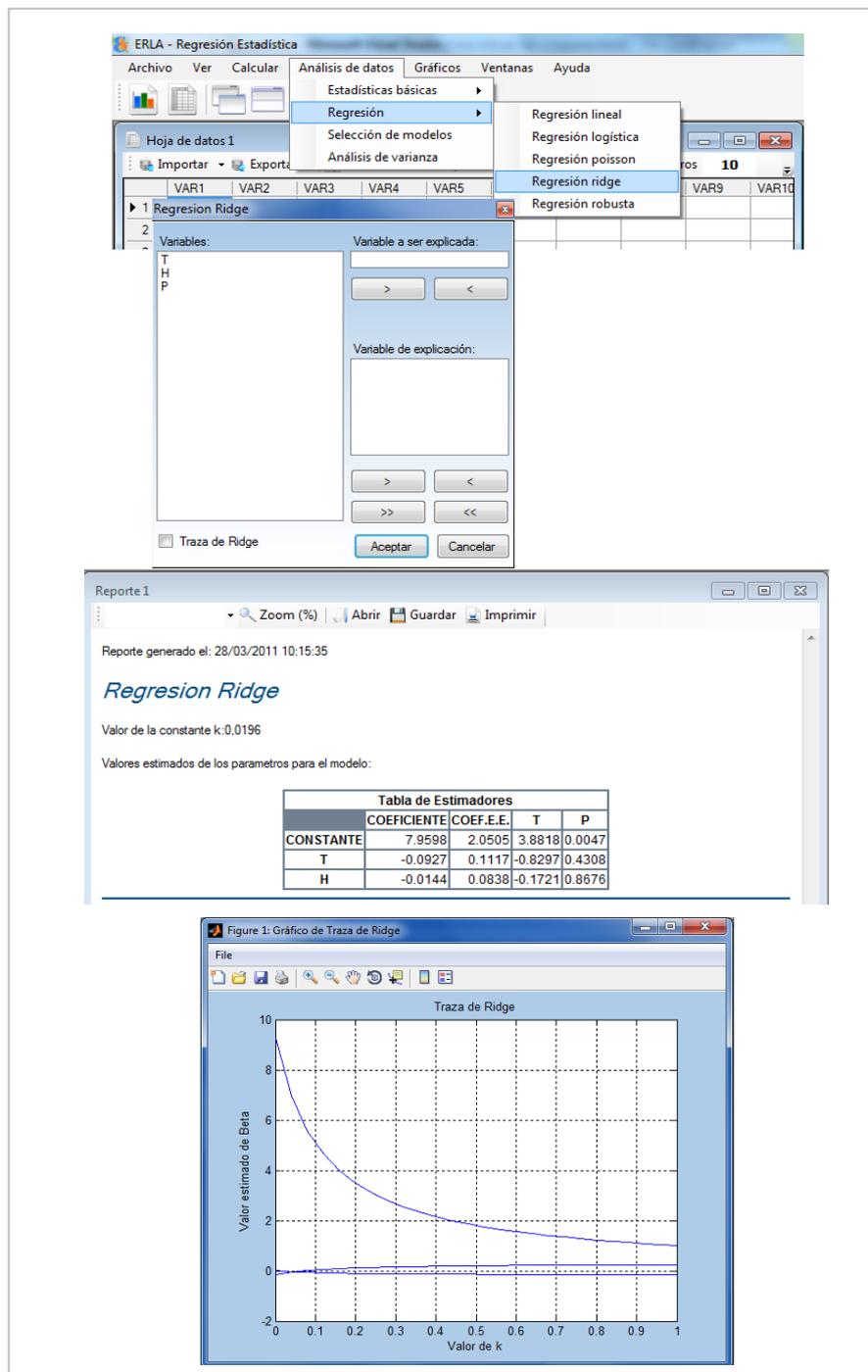
TEMPERATURA	25	27	29	31	32	33	34	35	36	30	24
HUMEDAD	17	18	19	22	25	27	29	33	23	20	15
PRODUCCIÓN	5.3	5.2	5.1	5	4.8	4.9	5.3	3.2	3.1	5.9	5.2

Para realizar la regresión Ridge en ERLA siga la siguiente secuencia de pasos:

1. **Barra de menú** > **Análisis de datos** > **Regresión** > **Regresión Ridge**
2. Seleccione las variables que desee y luego pulse el botón **Aceptar**.

Los resultados se presentan en la **ventana de reporte** junto con el gráfico correspondiente. Véase Figura 4.25.

Figura 4.25
ERLA 2011 Estadística ICM/ESPOL
"Regresión Ridge"



**Interpretación
de resultados**

Los resultados que se presentan en la **ventana de reporte** corresponden a:

- Valor de la Constante k es decir que del intervalo comprendido entre $[0 \ 1]$ el valor de 0.0196 hace que los coeficientes de regresión que tenían un valor demasiado grande empiezan a tener valores razonables.
- Tabla que contiene los valores estimados de los parámetros para el modelo, aquí tenemos los coeficientes, los coeficientes del Error Estándar, el Estadístico de Prueba F y el valor p .

En la **ventana de gráfico** se muestra la Traza de Ridge para los diferentes valores de k con respecto a los diferentes valores de $Beta$.

Regresión robusta

Ahora en el caso de regresión robusta se emplea el mismo ejemplo usado en regresión lineal simple, donde se trata de explicar la variable *P10* (*Identifico a los estudiantes de la ESPOL por su honestidad*) en función de la variable *P9* (*Identifico a los estudiantes de la ESPOL por su responsabilidad*).

Para dar una mejor ilustración de nuestro problema hemos contaminado el 20% de la muestra. Ahora mediante la sentencia:

1. Barra de menú > **Análisis de datos** > **Regresión** > **Regresión Ridge**
2. Seleccione las variables que desee y luego pulse el botón **Aceptar**. Véase Figura 4.26.

Los resultados se presentan en la **ventana de reporte** junto con el gráfico correspondiente. Véase Figura 4.27.

Figura 4.26
ERLA 2011 Estadística ICM/ESPOL
"Regresión robusta"

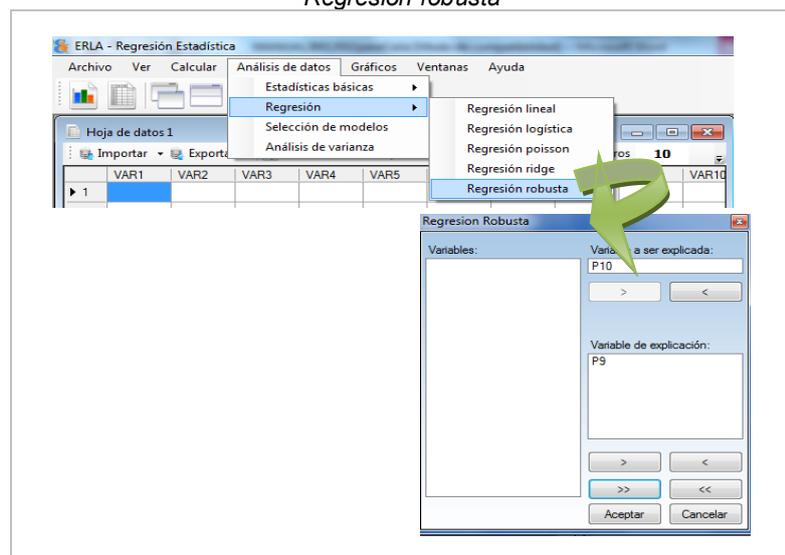
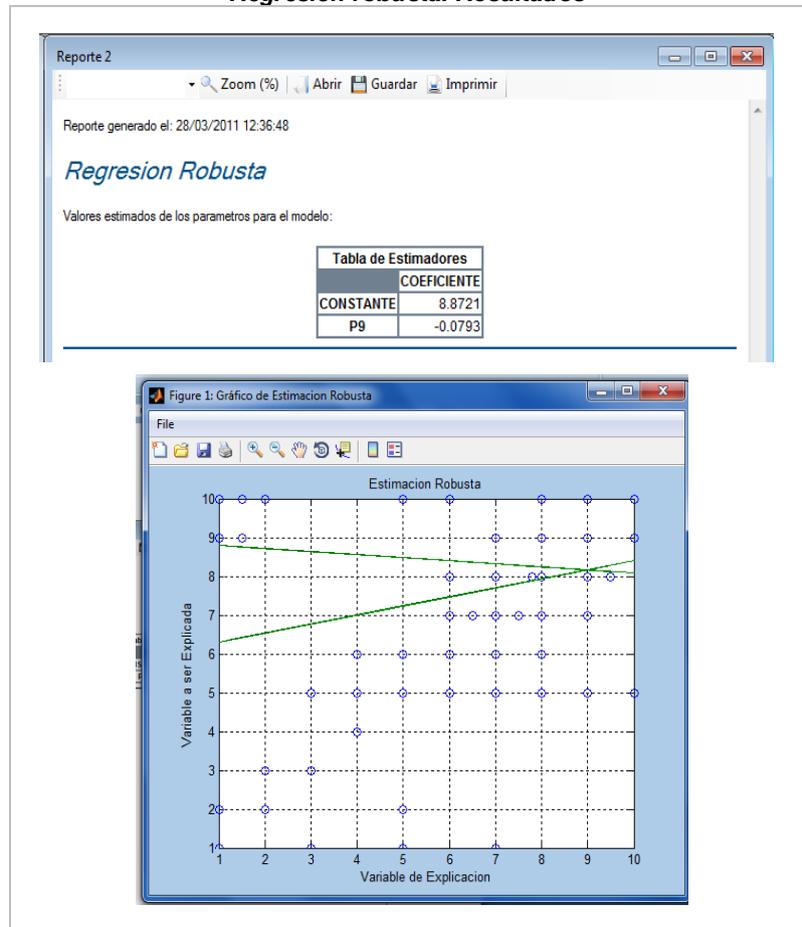


Figura 4.27
ERLA 2011 Estadística ICM/ESPOL
“Regresión robusta: Resultados”

Salida en
 ventana de
 reporte

Salida en
 ventana de
 gráfico



**Interpretación
 de resultados**

En la **ventana de reporte** se presenta la tabla que contiene los coeficientes de los parámetros para el modelo.

En la **ventana de gráfico** se muestra la representación gráfica de la estimación robusta, que corresponde al diagrama de dispersión de los datos con las rectas de Estimación Robusta y de Mínimos Cuadrados. Observamos que la estimación robusta proporciona un mejor ajuste a la tendencia de las observaciones mientras que la recta de estimación por Mínimos Cuadrados no se ajusta, debido a la presencia de valores aberrantes.

Análisis de varianza

La técnica del análisis de varianza en ERLA se realiza a través de la siguiente secuencia:

Barra de menú > Análisis de datos > Análisis de varianza

ERLA permite al usuario realizar análisis de varianza de n factores. Las instrucciones para ello se presentan a continuación.

Análisis de varianza de un factor

Para poder ilustrar el uso de esta técnica estadística vamos a utilizar los datos de la Tabla 4.2, que se refieren al nivel de producción lechera observado en 12 vacas durante 3 meses después de haberles aplicado tres niveles distintos de levadura en su ración alimenticia.

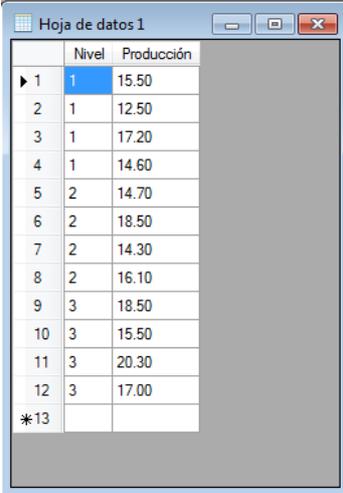
Tabla 4.2
ERLA 2011 Estadística ICM/ESPOL
"Datos para el análisis de varianza de una vía: Producción lechera"

Niveles del factor	Producción lechera			
1	15.5	12.5	17.2	14.6
2	14.7	18.5	14.3	16.1
3	18.5	15.5	20.3	17.0

En el ejemplo propuesto se distinguen 3 niveles del factor (nivel de levadura), y por cada nivel se han tomado 4 observaciones. Para poder realizar el análisis de varianza de un factor del ejemplo propuesto, los datos se deben ingresar especificando el nivel al que pertenece cada valor como se indica en la Figura 4.28.

Figura 4.28

ERLA 2011 Estadística ICM/ESPOL
"Ingreso de datos para el análisis de varianza de un factor"



The screenshot shows a window titled "Hoja de datos 1" containing a data table. The table has two columns: "Nivel" and "Producción". The data is as follows:

	Nivel	Producción
▶ 1	1	15.50
2	1	12.50
3	1	17.20
4	1	14.60
5	2	14.70
6	2	18.50
7	2	14.30
8	2	16.10
9	3	18.50
10	3	15.50
11	3	20.30
12	3	17.00
*13		

Como se trata de un análisis de un solo factor, se tienen dos columnas la primera en este caso para identificar el nivel del factor y la segunda para los valores de la variable observada. Para realizar el análisis de varianza se debe indicar en ERLA, cuál es la variable de observación y cuál es el factor. Véase Figura 4.29.

Figura 4.29
ERLA 2011 Estadística ICM/ESPOL
“Análisis de varianza de un factor”

The screenshot displays the ERLA 2011 software interface. The 'ANOVAs n-vías' dialog box is open, with 'Respuesta' selected as the 'Variable de medición' and 'Factor' as the 'Factores'. The 'Interacción entre factores' checkbox is unchecked. The 'Aceptar' button is highlighted with a green arrow.

The main window shows a report titled 'Análisis de Varianza de 1 Vía'. The report includes the following data:

Reporte generado el: 28/03/2011 11:21:37

Variable de medición: Respuesta $s = 1.9708$ $R^2 = 0.3293$

ANOVA					
	G.L.	S.C.	M.C.	F	P
FACTOR	2	17.185	8.5825	2.2096	0.1657
ERROR	9	34.9575	3.8842		
TOTAL	11	52.1225			

Tablas de Coeficientes del Modelo

COEFICIENTES	
NOMBRE	VALOR
Constante	16.225
Factor=1	-1.275
Factor=2	-0.325
Factor=3	1.6

Intervalos de Confianza para la media de los tratamientos

95% de Confianza				
TRATAMIENTO	NO. DE OBS.	L. INFERIOR	MEDIA	L. SUPERIOR
1	4	12.7208	14.95	17.1792
2	4	13.6708	15.9	18.1292
3	4	15.5958	17.825	20.0542

Intervalos de Confianza para la diferencia de medias entre los tratamientos

Intervalos de Confianza para la diferencia de medias entre los tratamientos

95% de Confianza				
TRATAMIENTO 1	TRATAMIENTO 2	L. INFERIOR	DIF. DE MED.	L. SUPERIOR
1	2	-4.8409	-0.95	2.9409
1	3	-6.7659	-2.875	1.0159
2	3	-5.8159	-1.925	1.9659

Minima diferencia significativa de Fisher

LSD de Fisher				
TRATAMIENTO 1	TRATAMIENTO 2	V.A.D.M.	t	p
1	2	0.95	0.6817	0.2563
1	3	2.875	2.063	0.0346
2	3	1.925	1.3813	0.1003

En la **ventana de reporte** se presenta:

- El valor de s es decir el estimador de la varianza del modelo, que como ya sabemos es homocedástica (la misma en todo el modelo). En este caso se obtuvo el valor de $s=1,9708$. También tenemos el R^2 , que es la proporción de variabilidad en los datos explicada por el modelo de análisis de varianza, para este ejemplo se tiene que el *Nivel de levadura* explica cerca de 32,93% de la variabilidad.
- Tabla ANOVA con todos sus valores característicos: Fuentes de variación (1era columna), grados de libertad (*G.L.*), sumas y medias cuadráticas (*S.C.* y *M.C.*, respectivamente), el estadístico de prueba F (F) y el valor p (P).
- Tabla de coeficientes del modelo. Aquí tenemos el valor de la media general del modelo (μ), y después los valores de los efectos de cada tratamiento.
- Intervalos de confianza para las medias de los tratamientos utilizando un nivel de confianza del 95%. Se puede distinguir en los resultados de la tabla el límite inferior (L. INFERIOR) y el límite superior (L. SUPERIOR) del intervalo para la media de cada tratamiento.
- Intervalos de confianza para la diferencia de medias, utilizando un nivel de confianza del 95%. Se consideran todas las posibles combinaciones de las medias. Se puede distinguir en los resultados de la tabla el límite inferior (L. INFERIOR) y el límite superior (L. SUPERIOR).
- Mínima diferencia significativa de Fisher, donde se muestra en la columna *VADM*, el valor absoluto de la diferencia de los pares de tratamientos, junto al valor del estadístico t para realizar la comparación correspondiente. Dos tratamientos son significativamente diferentes siempre que: $VADM > t$.

Análisis de varianza de dos o más factores

Ahora pasaremos a un diseño experimental en el que existan dos o más factores. Para ello se consideran los datos de la Tabla 4.3, los cuales corresponden al experimento de observar el tiempo de vida efectiva de una batería sometiéndola a diferentes temperaturas y tipos de materiales de la placa. Se han considerado tres niveles de temperatura ($15^{\circ}F$, $70^{\circ}F$, $125^{\circ}F$) de acuerdo al medio ambiente donde se usara el producto, y luego hay tres tipos de materiales para el diseño de la placa. Se han probado cuatro baterías con cada combinación de material y temperatura de forma aleatoria.

Tabla 4.3
ERLA 2011 Estadística ICM/ESPOL
"Datos para el análisis de varianza de dos factores de la vida (en horas) del diseño de la batería"

		Temperatura (F)		
		15	70	125
Tipo de material	1	130; 155; 74; 180	34; 40; 80; 75	20; 70; 82; 58
	2	150; 188; 159; 126	136; 122; 106; 115	25; 70; 58; 45
	3	138; 110; 168; 160	174; 120; 150; 139	96; 104; 82; 60

En este ejemplo se distinguen 3 niveles del primer factor (tipo de material) y 3 niveles del segundo factor (temperatura). Para poder realizar el análisis de varianza de dos vías, los datos deben ingresarse tal y como se muestra en la Figura 4.30.

Para el caso de dos factores, se tienen 3 columnas, la primera corresponderá al primer factor: *Tipo de material*, el Segundo Factor: *Temperatura* y la tercera, a la variable de observación, en este caso *Vida efectiva de la batería*. Para el caso de n factores se tendrán n+1 columnas, las primeras n correspondientes a los n factores y la última correspondiente a la variable de observación. Para obtener los resultados el usuario debe indicar en ERLA cuáles son los factores y cuál es la variable de observación o medición tal como se puede ver en la Figura 4.31.

Figura 4.30
 ERLA 2011 Estadística ICM/ESPOL
 “Ingreso de datos para el análisis de varianza de dos o más factores”

	Tipo de material	Temperatura	Tiempo/vida/diseño
15	2,00	1,00	159,00
16	2,00	1,00	126,00
17	2,00	2,00	136,00
18	2,00	2,00	122,00
19	2,00	2,00	106,00
20	2,00	2,00	115,00
21	2,00	3,00	25,00
22	2,00	3,00	70,00
23	2,00	3,00	58,00
24	2,00	3,00	45,00
25	3,00	1,00	138,00
26	3,00	1,00	110,00
27	3,00	1,00	168,00
28	3,00	1,00	160,00
29	3,00	2,00	174,00
30	3,00	2,00	120,00
31	3,00	2,00	150,00
32	3,00	2,00	139,00
33	3,00	3,00	96,00
34	3,00	3,00	104,00
35	3,00	3,00	82,00
36	3,00	3,00	60,00
►*37			

Figura 4.31
ERLA 2011 Estadística ICM/ESPOL
“Análisis de varianza de dos o más factores”

ANOVA n-vías

Variables:
 Variable de medición: Tiempo/vida/diseño

Factores:
 Tipo de material
 Temperatura

Interacción entre factores

Análisis de Varianza de 2 Vías

Modelo sin interacciones

Variable analizada: Tiempo/vida/diseño $s = 29.9702$ $R^2 = 0.6414$

ANOVA					
	G.L.	S.C.	M.C.	F	P
Tipo de material	2	10683.7	5341.86	5.9472	0.0065
Temperatura	2	39118.7	19559.4	21.7759	0
ERROR	31	27844.5	898.211		
TOTAL	35	77647			

Tablas de Coeficientes del Modelo

COEFICIENTES	
NOMBRE	VALOR
Constante	105.528
Tipo de material=1	-22.3611
Tipo de material=2	2.8056
Tipo de material=3	19.5556
Temperatura=1	39.3056
Temperatura=2	2.0556
Temperatura=3	-41.3611

Si necesario, se puede incluir las interacciones entre factores con tan solo seleccionar la opción **Interacción entre factores**. Los resultados serán los siguientes:

- El valor de s es decir el estimador de la varianza del modelo, que como ya sabemos es homocedástica (la misma en todo el modelo). En este caso se obtuvo el valor de $s=29,9702$. También tenemos el R^2 , que es la proporción de variabilidad en los datos explicada por el modelo de análisis de varianza, para este ejemplo se tiene que el modelo explica cerca de 64,14% de la variabilidad.
- Tabla ANOVA con todos sus valores característicos: Fuentes de Variación (1era columna), aquí se incluyen los dos factores del experimento, es decir *Tipos de material* y *Temperatura*; grados de libertad (*G.L.*), sumas y medias cuadráticas (*S.C.* y *M.C.*, respectivamente), el estadístico de prueba F (*F*) y el valor p (*P*).
- Tabla de coeficientes del modelo. Aquí tenemos el valor de la media general del modelo (μ), y después los valores de los efectos de los diferentes niveles de cada factor.



RECOMENDACIONES

Al momento de usar ERLA, debe considerar los siguientes puntos:

- Los cálculos son realizados en MATLAB, y por ello las expresiones ingresadas en la **Calculadora** son expresadas empleando la misma sintaxis de MATLAB. En caso de que las expresiones no sean las correctas se mostrará un mensaje de error que proviene del compilador de MATLAB.
- El número de registros en cada una de las variables de una misma **ventana de datos** debe ser el mismo. Si necesita ingresar variables con número de registros distintos agregue una nueva **ventana de datos** por cada variable. Para realizar los cálculos correspondientes ERLA toma en cuenta todos los registros de la **ventana de datos**, los registros con valores faltantes no serán considerados.
- Debido a que ERLA utiliza el MCR (MATLAB Compiler Runtime) de MATLAB, el tiempo que tarda en iniciar es el mismo tiempo que tarda en iniciar MATLAB.
- ERLA no tiene un número limitado de registros para trabajar, sus limitaciones provienen de las características de hardware del equipo donde se instala.
- Se debe tener en cuenta que ERLA es un software que se mantiene en desarrollo para mejorar sus versiones posteriores o agregar más técnicas estadísticas propuestas por los usuarios de esta aplicación. Para formar parte del proyecto visite: <http://code.google.com/p/erla>



REFERENCIAS BIBLIOGRÁFICAS

MINITAB. (2010), “*Meet Minitab 16*”, Quality Companion by Minitab y Quality Trainer by Minitab, Estados Unidos.

RODRÍGUEZ, L. (2007), “*MATLAB Conceptos básicos y programación*”, **Escuela Superior Politécnica del Litoral, Guayaquil – Ecuador.**

ANEXO I

Centro de Estudios e Investigaciones Estadísticas
PROYECTO DE MUESTREO Y DISEÑO DE CUESTIONARIOS
Imagen de las Universidades ecuatorianas en Guayaquil

Formulario a ser administrado a los habitantes de Guayaquil

1. CARACTERÍSTICAS GENERALES DEL INFORMANTE

1. Género: Masculino Femenino

2. Año de Nacimiento:

3. Estado Civil: Soltero Casado Divorciado Viudo Unión Libre

4. Nivel de Instrucción: Primario Secundario Superior Ninguno

Parroquia: _____
Zona: _____
Sector: _____

2. INSTITUCIONES

Instrucción: Para responder las preguntas 1 y 2 elija una opción de las proporcionadas por el entrevistador.

1. A su entender, ¿cuál es la Institución que mayor confianza inspira a los ecuatorianos?

2. ¿Qué Universidad considera usted es la mejor del Ecuador?

3. En una escala de 1 a 10, asigne una calificación a las siguientes Universidades de acuerdo a lo que usted conoce de ellas:

Universidad	Calificación	No sé	Universidad	Calificación	No sé
1. Escuela Politécnica Nacional (EPN)	<input type="checkbox"/>	<input type="checkbox"/>	8. Universidad de Especialidades Espiritu Santo (UEES)	<input type="checkbox"/>	<input type="checkbox"/>
2. Escuela Superior Politécnica del Ejército (ESPE)	<input type="checkbox"/>	<input type="checkbox"/>	9. Universidad de Guayaquil (Estatal)	<input type="checkbox"/>	<input type="checkbox"/>
3. Escuela Superior Politécnica del Litoral (ESPOL)	<input type="checkbox"/>	<input type="checkbox"/>	10. Universidad Politécnica Salesiana (UPS)	<input type="checkbox"/>	<input type="checkbox"/>
4. Universidad Católica de Guayaquil (UCSG)	<input type="checkbox"/>	<input type="checkbox"/>	11. Universidad San Francisco de Quito (USFQ)	<input type="checkbox"/>	<input type="checkbox"/>
5. Universidad Católica de Quito	<input type="checkbox"/>	<input type="checkbox"/>	12. Universidad Técnica Particular de Loja (UTPL)	<input type="checkbox"/>	<input type="checkbox"/>
6. Universidad Central del Ecuador (Quito)	<input type="checkbox"/>	<input type="checkbox"/>	13. Universidad Técnica de Ambato (UTA)	<input type="checkbox"/>	<input type="checkbox"/>
7. Universidad de Cuenca	<input type="checkbox"/>	<input type="checkbox"/>	14. Otra _____ (especifique)	<input type="checkbox"/>	<input type="checkbox"/>

3. ACERCA DE LA ESPOL

1. ¿Ha visitado alguna vez el Campus Gustavo Galindo (Prosperina) de la ESPOL? Sí No

2. ¿Tiene usted alguna relación con la ESPOL?

- Profesor Ex-profesor Tiene algún familiar relacionado con la ESPOL
 Estudiante Ex-estudiante Ninguna
 Empleado Ex-empleado Otra _____
(especifique)

Continúa...

... viene

3. ¿Sabe usted cuál es el animal que aparece en el logo de la ESPOL? Sí _____ No
(especifique)

4. ¿Conoce usted el nombre del Rector de la ESPOL?
 Nombre y Apellido Sólo el nombre Sólo el apellido No sé

5. ¿Sabe usted cómo acceder desde su hogar al campus Politécnico Gustavo Galindo? Sí No

Instrucción: Lea detenidamente las siguientes proposiciones y asigne a cada una de ellas una calificación entre uno (1) y diez (10). (Puede utilizar hasta dos decimales)

4. PROPOSICIONES	Calificación
1. Conozco información de las opciones profesionales que brinda la ESPOL.	
2. Me encuentro adecuadamente informado sobre asuntos relacionados con la ESPOL.	
3. Lo que caracteriza a la enseñanza en la ESPOL es la excelencia.	
4. La forma en que los bachilleres ingresan a la ESPOL se distingue por el grado de dificultad de los exámenes.	
5. La forma en que los bachilleres ingresan a la ESPOL se caracteriza por la seriedad con la que se examina a los estudiantes.	
6. Los estudios en la ESPOL se caracterizan por un alto contenido en <i>Ciencias</i> .	
7. Los estudios en la ESPOL se caracterizan por un alto contenido en <i>Informática</i> .	
8. Los estudios en la ESPOL se caracterizan por un alto contenido <i>Técnico</i> .	
9. Identifico a los estudiantes de la ESPOL por su responsabilidad.	
10. Identifico a los estudiantes de la ESPOL por su honestidad.	
11. ESPOL cumple con su misión de formar <i>Profesionales</i> de excelencia.	
12. La capacitación que se imparte en la ESPOL permite un adecuado desempeño a nivel <i>Profesional</i> .	
13. Ser graduado en la ESPOL da ventajas competitivas frente a sus potenciales empleadores.	
14. Ser graduado en la ESPOL concede ventajas frente a sus competidores de otras universidades.	
15. El trabajo que realiza la ESPOL tiene un impacto positivo en la <i>Comunidad</i> .	
16. La ESPOL se distingue por ser una Institución que efectúa <i>Investigación Científica</i> .	
17. Los <i>programas culturales</i> que ofrece la ESPOL en el campus Las Peñas son de alta calidad.	
18. Es notoria la preocupación que la ESPOL muestra con respecto al <i>Medio Ambiente</i> .	
19. La ESPOL es una institución en la que se evidencia la no relación con <i>Partido Político</i> alguno.	

Fecha en que completa el presente Formulario: ___/___/___ día/mes/año