GUIÓN 2 de PRÁCTICAS con Rcmdr Análisis bivariante de datos

Tablas de Contingencia

Se toman los siguientes datos correspondientes a 10 personas en los que se consideran el sexo y el color de ojos:

Sexo	Col.Ojos
mujer	negro
hombre	negro
hombre	azul
hombre	verde
hombre	verde
mujer	verde
mujer	negro
hombre	verde
mujer	azul
mujer	azul

Introducir la tabla anterior en RCommander

Método 1: Escribir la tabla en el Bloc de Notas, guardarlo e importarlo con Rcmdr.

Método 2: Escribir directamente en la Ventana de Instrucciones

```
sexo <-c("mujer","hombre","hombre","hombre","hombre",
"mujer","mujer","hombre","mujer","mujer")
ojos <-c("negro","negro","azul","verde","verde",
"verde","negro","verde","azul","azul")
OjoSexo <- data.frame(sexo,ojos)
Ejecutar
```

Ir a

Estadísticos \rightarrow Tablas de contingencia \rightarrow Tablas de doble entrada





Obsérvese, en la tabla de datos,

- ¿Cuántas de las personas son mujeres con ojos negros? (2)
- ¿Qué porcentaje son hombres con ojos azules? (10%)
- ¿Qué porcentaje son hombres? (50%)
- ¿Qué porcentaje tiene ojos verdes? (40%)

Diagramas de barras segmentados por variables

Cargamos la librería *lattice*, y después se dibuja un diagrama de barras agrupados según la variable sexo



Análisis de pares de variables cuantitativas

Por ejemplo, cargamos los datos de Anscombe

74 R Commander	
Fichero Editar Datos Estadísticos Gráficas Modelos Distribuciones Herramientas Ayuda	
R Conjunto de Mateixa Modelos	nodelo activo >
Cargar conjunto de datos	
Ventana de inst Fusionar conjuntos de datos	
Importar datos	
Conjunto de datos en paquetes Lista de conjuntos de datos en paquetes	
Conjunto de datos activo	unto
Modificar variables del conjunto de datos activo 🕨	13
	~
	>
Ventana de resultados	Ejecutar
	~
> library(lattice)	
<pre>> barchart(OjoSexo,groups=OjoSexo\$sexo)</pre>	
	~
	>
Mensajes	
[1] NOTA: Versión de R Commander 1.5-5: Sun Feb 21 00:06:43 2010	
[2] NOTA: El conjunto de datos OjoSexo tiene 10 filas y 2 columnas.	
<	>

74 Leer datos desde paquet		
Paquete (doble clic para seleccionar) car datasets lattice MASS	Conjunto de datos (doble clic para seleccionar) Adler Angell <u>Anscombe</u> Baumann	
O Introducir el nombre del conjunto de datos: Anscombe		
Ayuda sobre ei conjunco de dacos seleccionado Aceptar Cancelar Ayuda		

Diagramas de puntos

Ir a

Gráficas → Diagramas de Dispersión

Rellenar así los siguientes campos:

74 Diagrama de dispersión	
variable x (elegir una) variable y (elegir una) education education income income urban urban young young	
Options	Dibujar parámetros
Identificar Observaciones	Dibujar los caracteres
Desplazar aleatoriamente la variable x 🔲 Desplazar aleatoriamente la variable y 🗍	Tamaño del punto
Log eje-x	Tamaño del texto (ejes)
Log eje-y T Cajas de dispersión marginales T	Tamaño del texto (etiquetas de los ejes)
Línea de mínimos cuadrados 🛛 🔽	
Línea suavizada 📃	
Amplitud para suavizado 50	
Etiqueta del eje x Etiqueta del eje y	
<auto> <auto></auto></auto>	
X	>
Expresión de selección	
<todos casos="" los="" td="" válidos<=""><td></td></todos>	
Gráfica por grupos	
Aceptar Cancelar Ayuda	

Se obtiene el siguiente resultado:



Se pueden obtener todas las gráficas de dispersión al mismo tiempo:

Gráficas → Matriz de Diagramas de Dispersión

74 R Commander	
Fichero Editar Datos Estadísticos Gráficas Modelos Distribuciones Herramientas Ayuda	
Red Conjunto de datos: Anscombe Editar conjunto de datos Visualizar conjunto de datos Modelo: <a car")<="" href="https://www.editar.com/www</td><td>tivo></td></tr><tr><td>data(Anscombe, package=" td=""><td>4</td>	4
scatterplot.matrix(~education+income+urban+young, reg.line=lm, smooth=FALSE, span=0.5, diagonal = 'boxplot', data=Anscombe)	
Matriz de diagramas de dispersión	2
income urban	>
Ventana de resultados Ejecutar	
<pre>> scatterplot.matrix smooth=FALSE, span=0</pre> Líneas suavizadas I Líneas suaviz	
Gráfica por grupos Aceptar Cancelar Ayuda	2
Mensajes	
Use 'scatterplotMatrix' instead. See help("Deprecated") and help("car-deprecated").	~
	2



Covarianzas y Correlaciones

Para calcular las correlaciones:

Estadísticos → Resúmenes → Matriz de Correlaciones

🌠 R Commander	×
Fichero Editar Datos Estadísticos Gráficas Modelos Distribuciones Herramientas Ayuda	
Rendr Conjunto de datos: Anscombe Editar conjunto de datos Visualizar conjunto de datos Modelo: <a)="" <="" car")="" cor(anscombe[,c("education","income","urban","young")],="" correlaciones="" de="" href="https://www.edu.com/workship-like/wor</td><td></td></tr><tr><td>Ventana de instrucciones</td><td></td></tr><tr><td><pre>data(Anscombe, package=" matriz="" pre="" use="complete.obs">	9
I Variables (elegir dos o más) education education income urban vounq Tipo de correlaciones Ventana de resultados Coeficiente de Pearson Coeficiente de Spearman Ejecutar Parcial Parcial p-valor pareado de las correlaciones de Pearson o Spearman	
<pre>Aceptar Cancelar Ayuda > cor(Anscombe[,c("education","income","urban","young")], use="complete.obs")</pre>	
education income urban young education 1.0000000 0.6675773 0.2633238 0.3114855 income 0.6675773 1.0000000 0.6854580 -0.1623600 urban 0.2633238 0.6854580 1.0000000 -0.1386334 young 0.3114855 -0.1623600 -0.1386334 1.0000000	
Mensaies	
 [1] NOTA: Versión de R Commander 1.5-5: Sun Feb 21 16:10:08 2010 [2] NOTA: El conjunto de datos Anscombe tiene 51 filas y 4 columnas. 	
×	

La matriz de correlaciones es **simétrica** y tiene un 1 **siempre** en la diagonal

Para calcular las **covarianzas**:

En la Ventana de Instrucciones sustituir el comando cor por el comando cov

74 R Commander	
Fichero Editar Datos Estudísticos Gráficas Modelos Distribuciones Herramientas Ayuda	
R, Conjunto de datos: Anscombe Editar conjunto de datos Visualizar conjunto de datos Modelo: <no< td=""><td>hay modelo activo></td></no<>	hay modelo activo>
Condur · · · · · · · · · · · · · · · · · · ·	
data(lnscombe, nackage="car")	
<pre>cov(Anscombe[,c("education","income","urban","young")], use="complete</pre>	.obs")
4	
<	>
Ventana de resultados	Ejecutar
<pre>> cov(Anscombe[,c("education","income","urban","young")], use="comple</pre>	te.obs")
education income urban young	
education 2158.0196 17367.51 1851.3369 346.6984	
urban 1851.3369 58097.39 22905.2549 -502.7149	
young 346.6984 -2178.58 -502.7149 574.0804	
	~
3	>
Mensajes	
[1] NOTA: Versión de R Commander 1.5-5: Sun Feb 21 16:10:08 2010	
[2] NOTA: El conjunto de datos Anscombe tiene 51 filas y 4 columnas.	
	×
S.	7

Se obtiene una matriz **simétrica** con los valores de las **varianzas** de cada variable en la **diagonal**.

Regresión Lineal

Para calcular una regresión lineal:

Estadísticos \rightarrow Ajuste de Modelos \rightarrow Regresión lineal

74 R. Commander	
Fichero Editar Datos Estadísticos Gráficas Modelos Distribuciones Herramientas Ayuda	
Reg	gModel.1
Ventana de instrucciones	1
data(Anscombe, package="car")	<u></u>
Regnodel.1 <- Im(income~education, data=Anscompe)	
Regresión lineal	
Introducir un nombre para el modelo: RegModel.1 Variable explicada (elegir una) Variables explicativas (elegir una o más) education income urban young Expresión do selección	~
Ventana de resultados	Ejecutar
	~
Aceptar Cancelar Ayuda	
Call:	
lm(formula = income ~ education, data = Anscombe)	
Residuals: Min 1Q Median 3Q Max -821.25 -341.15 8.47 196.23 1057.37	
Coefficients:	
Estimate Std. Error t value Pr(> t)	
(Intercept) 1645.383 258.539 6.364 6.41e-08 ***	
education 8.048 1.282 6.276 8.76e-08 ***	
 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1	
Residual standard error: 421.2 on 49 degrees of freedom	
Multiple R-squared: 0.4457, Adjusted R-squared: 0.4343	
F-statistic: 39.39 on 1 and 49 DF, p-value: 8.762e-08	~
	>
	1771
 [1] NOTA: version de R Commander 1.5-5: Sun Feb 21 16:10:08 2010 [2] NOTA: El conjunto de datos Anscombe tiene 51 filas y 4 columnas. 	
<	>

El valor del punto de corte en la ordenada es a = 1645.383

El valor de la pendiente es b = 8.048

La desviación estándar residual es $\sqrt{421.2} = 20.52$

El coeficiente de determinación R^2 (cuadrado del coeficiente de correlación lineal) es 0.4457

Gráficas de residuos

Para observar las gráficas de residuos:

 $Modelos \rightarrow Gráficas \rightarrow Gráficas básicas de diagnósticos$

74 R Commander		
Fichero Editar Datos Estadísticos Gráficas Modelos Distribuciones Herramientas Ayuda		Taria da Anterio
Fichero Editar Datos Estadísticos Gráficas Modelos Distribuciones Herramientas Ayuda Selecciona el modelo activo Ventana de instrucciones data (Anacombe, package="car") Citar el conjunto de datos: Anacombe, package="car") Citar el conjunto de datos el normación Bayesiano (BIC) Selección de modelo paso a paso Filtrer el conjunto de datos editor Citar el conjunto de datos el normación Bayesiano (BIC) Selección de modelo paso a paso Filtrer el conjunto de datos activo Test de hipótesis Diagnósticos numéricos Gráficas	Model.1 Gráficas básicas de dia próstuco Gráfica de comparación de cuantiles de los residuos Gráfica de componentes+residuos Gráfica de variables agregadas Gráfica de influencia Gráfica de los efectos	Ejecular
Mensajes [1] NOTA: Versión de R Commander 1.5-5: Sun Feb 21 16:10:08 2010 [2] NOTA: El conjunto de datos Anscombe tiene 51 filas y 4 columnas		× ×

Obsérvense las dos primeras gráficas que aparecen en la salida.

Los residuos oscilan alrededor de la recta igual a 0 y no muestran *tendencias* \rightarrow La regresión lineal es **adecuada**.

Los residuos se distribuyen como una distribución normal.



Visualizaciones con TeachingDemos

Se carga primero el plug-in correspondiente a *TeachingDemos*:

% R Commander	
Fichero Editar Datos Estadísticos Gráficas Modelos Distribuciones	Herramientas Ayuda
Reference de la conjunto de datos activo> Editar con Ventana de instrucciones	Cargar paquete(s) Cargar plugn(c) de Rondr Opciones
<	
Ventana de resultados	Ejecutar
< Marania	>
rmensages [1] NOTA: Versión de R Commander 1.5-5: Sat F	'eb 27 16:23:09 2010
2	



Aparece un nuevo menú denominado Demos

Marcar y pinchar en los submenús Simple linear regression y Simple correlation:

Observar el efecto de cambiar el número y disposición de los puntos en ambos casos (estudiar la teoría del tema 3).









