

T2. Taller encuestas. Descriptivos 2*

VIII Jornadas R Albacete 2016**

Álvaro Hernández Vicente, Elvira Ferre Jaén, Antonio José Perán Orcajada, Ana Belén
Marín Valverde, Antonio Maurandi López***

17 de noviembre de 2016

Índice

1. Lectura de datos	1
2. Descriptivos con paquete tables	1
2.1. Variable actividadS	2
2.2. Variable ingresos	3
2.3. Variable nivelIngles	3
3. Descriptivos con paquete likert	5
3.1. Según origen	5
3.2. Según sexo	7
Referencias y bibliografía	9

1. Lectura de datos

Se leen los datos `saeraq.Rdata`, que provienen del fichero `raq.dat` utilizado en (A. Field, Miles, and Field 2012) (disponible en la web del libro).

```
load( "saeraq.RData" )
```

2. Descriptivos con paquete tables

Se realizan descriptivos de cada variable según `sexo` y `origen` con el paquete `tables` (Murdoch (2016)).

```
# library( tables )
opt <- booktabs() # booktabs para las tablas

# Función para imprimir tablas en latex
tablaLatex <- function( tabla, caption = NULL ){
  cat( '\\begin{table} \\centering\\n' )
  if( !is.null( caption ) ) cat( paste0( '\\caption{', caption, '}\\n' ) )
  latex( tt )
  cat( '\\end{table}' )
}
```

* `doc:T2_descriptivos2.Rmd`

** <http://r-es.org/8jornadasR/>

*** Servicio de Apoyo Estadístico; alvarohv@um.es, elvira@um.es, antoniojose.peran@um.es, anabelen.marin4@um.es, amaurandi@um.es



2.1. Variable actividadS

```
tt <- tabular( ( Sexo = sexo ) + ( Total = 1 ) ~ ( `Actividad S` = actividadS ) +
              ( Total = 1 ), data = df )

# html( tt, options = htmloptions( HTMLcaption = "Variable actividadS según sexo.", pad = TRUE ) )
tablaLatex( tt, caption = "Variable \\texttt{actividadS} según \\texttt{sexo}." )
```

Tabla 1: Variable actividadS según sexo.

Sexo	Actividad S				Total
	Nada	Poco	Mucho	Muchísimo	
Mujer	466	487	509	465	1927
Hombre	147	171	160	166	644
Total	613	658	669	631	2571

```
tt <- tabular( ( Origen = origen ) + ( Total = 1 ) ~ ( `Actividad S` = actividadS ) +
              ( Total = 1 ), data = df )

# html( tt, options = htmloptions( HTMLcaption = "Variable actividadS según origen.", pad = TRUE ) )
tablaLatex( tt, caption = "Variable \\texttt{actividadS} según \\texttt{origen}." )
```

Tabla 2: Variable actividadS según origen.

Origen	Actividad S				Total
	Nada	Poco	Mucho	Muchísimo	
Albacete	288	282	308	300	1178
Murcia	263	286	292	271	1112
Helsinki	62	90	69	60	281
Total	613	658	669	631	2571

```
tt <- tabular( ( Sexo = sexo ) * ( Origen = origen ) + ( Total = 1 ) ~
              ( `Actividad S` = actividadS ) + ( Total = 1 ), data = df )

# html( tt, options = htmloptions( HTMLcaption = "Variable actividadS según origen y sexo.", pad = TRUE ) )
tablaLatex( tt, caption = "Variable \\texttt{actividadS} según \\texttt{origen} y \\texttt{sexo}." )
```

Tabla 3: Variable actividadS según origen y sexo.

Sexo	Origen	Actividad S				Total
		Nada	Poco	Mucho	Muchísimo	
Mujer	Albacete	212	206	238	221	877
	Murcia	205	210	220	197	832
	Helsinki	49	71	51	47	218
Hombre	Albacete	76	76	70	79	301
	Murcia	58	76	72	74	280
	Helsinki	13	19	18	13	63
Total		613	658	669	631	2571

2.2. Variable ingresos

```
ic1 <- function(x){
  mean( x ) - qt( 0.975, df = length( x ) - 1 ) * sd( x ) / sqrt( length( x ) )
}

ic2 <- function(x){
  mean( x ) + qt( 0.975, df = length( x ) - 1 ) * sd( x ) / sqrt( length( x ) )
}

tt <- tabular( ( Sexo = sexo ) + ( Origen = origen ) ~ ( Ingresos = ingresos ) *
              ( ic1 + mean + sd + ic2 + median ) + ( Total = 1 ), data = df )

# html( tt, options = htmloptions( HTMLcaption = "Variable ingresos según origen y sexo.", pad = TRUE )
tablaLatex( tt, caption = "Variable \\texttt{ingresos} según \\texttt{origen} y \\texttt{sexo}." )
```

Tabla 4: Variable ingresos según origen y sexo.

		Ingresos					Total
		ic1	mean	sd	ic2	median	
Sexo	Mujer	29262	29670	9129	30078	25906	1927
	Hombre	28982	29679	9000	30375	31628	644
Origen	Albacete	28917	29435	9059	29953	24651	1178
	Murcia	29311	29849	9143	30387	32055	1112
	Helsinki	28901	29965	9066	31030	32088	281

```
tt <- tabular( ( Sexo = sexo ) * ( Origen = origen ) + ( Total = 1 ) ~
              ( Ingresos = ingresos ) * ( ic1 + mean + sd + ic2 + median ) +
              ( Total = 1 ), data = df )

# html( tt, options = htmloptions( HTMLcaption = "Variable ingresos según origen y sexo.", pad = TRUE )
tablaLatex( tt, caption = "Variable \\texttt{ingresos} según \\texttt{origen} y \\texttt{sexo}." )
```

Tabla 5: Variable ingresos según origen y sexo.

		Ingresos					Total
Sexo	Origen	ic1	mean	sd	ic2	median	
Mujer	Albacete	28667	29267	9046	29866	24046	877
	Murcia	29181	29807	9201	30433	31833	832
	Helsinki	29552	30769	9121	31987	34367	218
Hombre	Albacete	28894	29925	9094	30957	32715	301
	Murcia	28919	29975	8983	31032	33033	280
	Helsinki	25078	27182	8357	29287	23057	63
Total		29320	29672	9095	30024	27771	2571

2.3. Variable nivelInglés

```
tt <- tabular( ( Sexo = sexo ) + ( Total = 1 ) ~ ( `Nivel de inglés` = nivelInglés ) +
              ( Total = 1 ), data = df )
```

```
# html( tt, options = htmloptions( HTMLcaption = "Variable nivelIngles según sexo.", pad = TRUE ) )
tablaLatex( tt, caption = "Variable \\texttt{nivelIngles} según \\texttt{sexo}." )
```

Tabla 6: Variable nivelIngles según sexo.

Sexo	Nivel de inglés										Total
	Nulo	CasiNulo	A1	A2	B1	B2	C1	C2	IsabelII	Shakespeare	
Mujer	227	210	343	345	354	278	109	56	3	2	1927
Hombre	70	79	108	110	119	99	36	23	0	0	644
Total	297	289	451	455	473	377	145	79	3	2	2571

```
tt <- tabular( ( Origen = origen ) + ( Total = 1 ) ~ ( `Nivel de inglés` = nivelIngles ) +
( Total = 1 ), data = df )
```

```
# html( tt, options = htmloptions( HTMLcaption = "Variable nivelIngles según origen.", pad = TRUE ) )
tablaLatex( tt, caption = "Variable \\texttt{nivelIngles} según \\texttt{origen}." )
```

Tabla 7: Variable nivelIngles según origen.

Origen	Nivel de inglés										Total
	Nulo	CasiNulo	A1	A2	B1	B2	C1	C2	IsabelII	Shakespeare	
Albacete	130	130	204	212	208	180	71	43	0	0	1178
Murcia	132	130	194	187	215	157	63	29	3	2	1112
Helsinki	35	29	53	56	50	40	11	7	0	0	281
Total	297	289	451	455	473	377	145	79	3	2	2571

```
tt <- tabular( (Sexo = sexo ) * ( Origen = origen ) + ( Total = 1 ) ~
( `Nivel de inglés` = nivelIngles ) + ( Total = 1 ), data = df )
```

```
# html( tt, options = htmloptions( HTMLcaption = "Variable nivelIngles según origen.", pad = TRUE ) )
tablaLatex( tt, caption = "Variable \\texttt{nivelIngles} según \\texttt{origen}." )
```

Tabla 8: Variable nivelIngles según origen.

Sexo	Origen	Nivel de inglés										Total
		Nulo	CasiNulo	A1	A2	B1	B2	C1	C2	IsabelII	Shakespeare	
Mujer	Albacete	92	101	159	159	160	126	50	30	0	0	877
	Murcia	106	89	150	139	152	121	50	20	3	2	832
	Helsinki	29	20	34	47	42	31	9	6	0	0	218
Hombre	Albacete	38	29	45	53	48	54	21	13	0	0	301
	Murcia	26	41	44	48	63	36	13	9	0	0	280
	Helsinki	6	9	19	9	8	9	2	1	0	0	63
Total		297	289	451	455	473	377	145	79	3	2	2571



3. Descriptivos con paquete likert

Se realizan gráficos descriptivos de los datos agrupando por `origen` o `sexo` con el paquete `likert` (Bryer and Speersschneider (2015)).

```
# library( likert )
dfLikert <- df[ , grep( "^Q", colnames( df ) ) ]
colnames( dfLikert ) <- dicc[ grep( "^Q", dicc$item ), "spanish" ]

bloque1 <- 1:8
bloque2 <- 9:17
bloque3 <- 18:23

items01 <- likert( items = dfLikert[ , bloque1 ], grouping = df$origen )
items02 <- likert( items = dfLikert[ , bloque2 ], grouping = df$origen )
items03 <- likert( items = dfLikert[ , bloque3 ], grouping = df$origen )

itemsS1 <- likert( items = dfLikert[ , bloque1 ], grouping = df$sexo )
itemsS2 <- likert( items = dfLikert[ , bloque2 ], grouping = df$sexo )
itemsS3 <- likert( items = dfLikert[ , bloque3 ], grouping = df$sexo )
```

3.1. Según origen

```
plot( items01 )
```

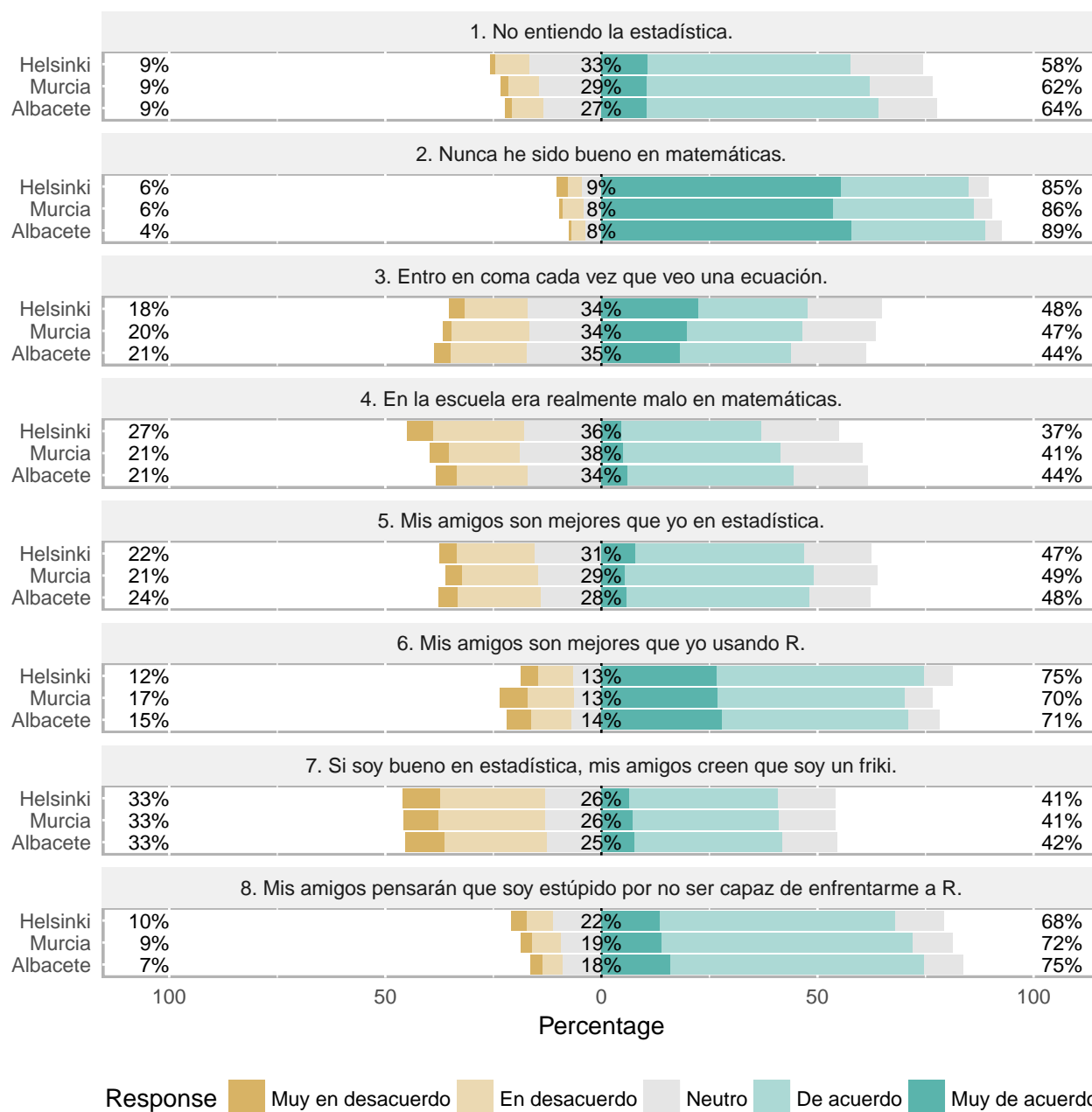


Figura 1: Gráfico de barras según origen de los ítems 1 a 8.

```
plot( items01, type = "density" )
```

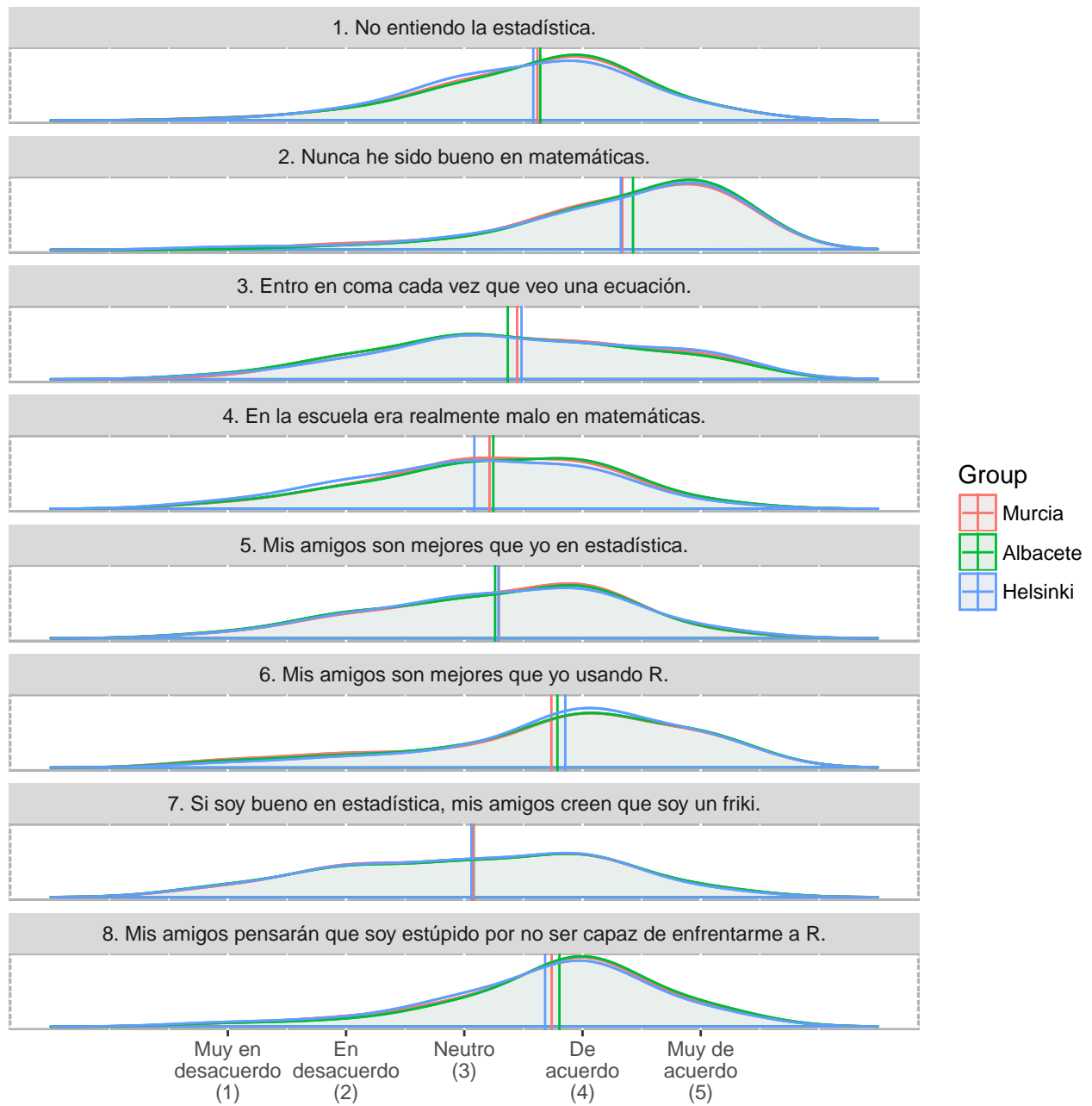


Figura 2: Gráfico de densidad según origen de los ítems 1 a 8.

3.2. Según sexo

```
plot( itemsS1 )
```

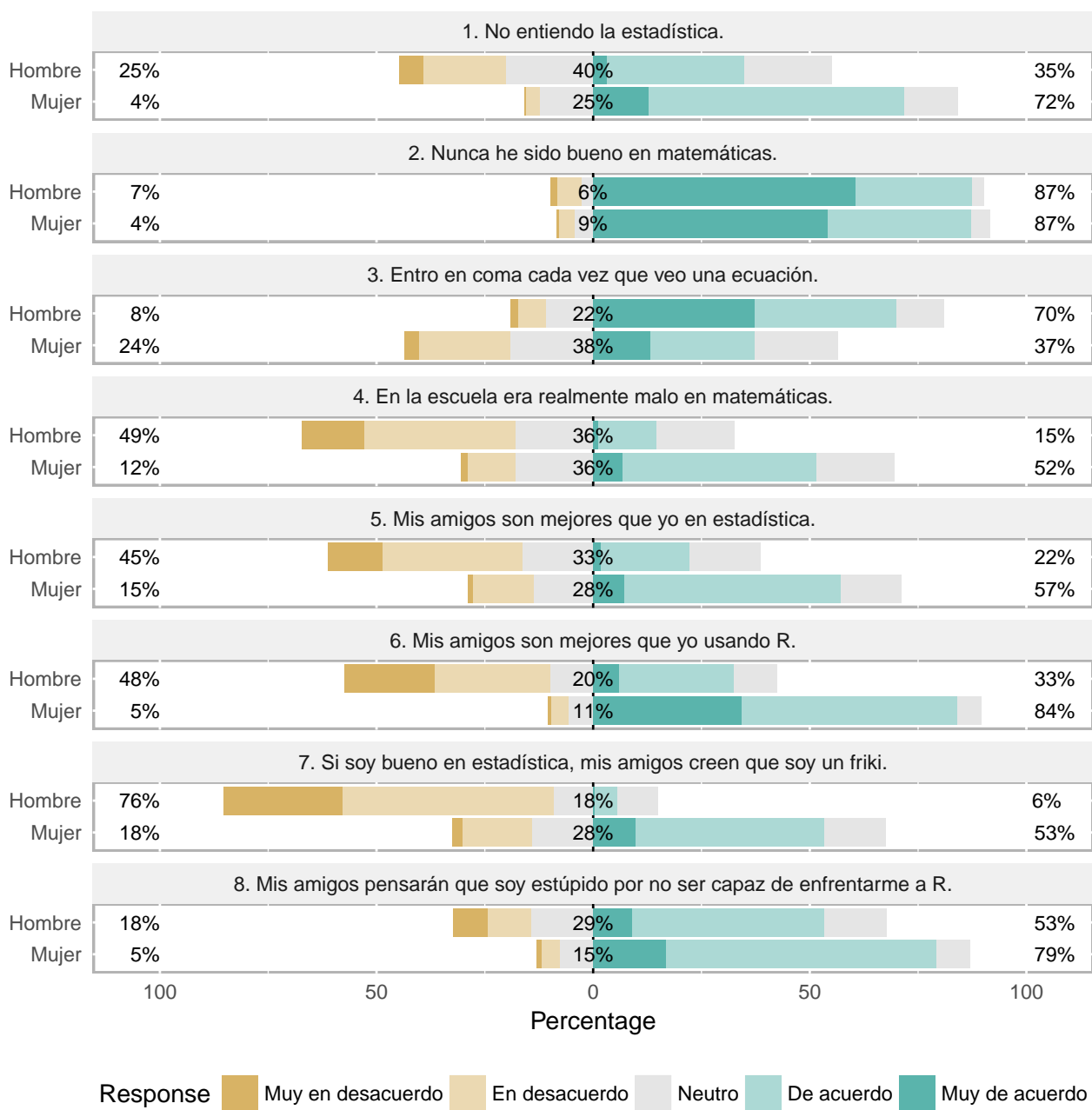


Figura 3: Gráfico de barras según sexo de los ítems 1 a 8.

```
plot( itemsS1, type = "density" )
```

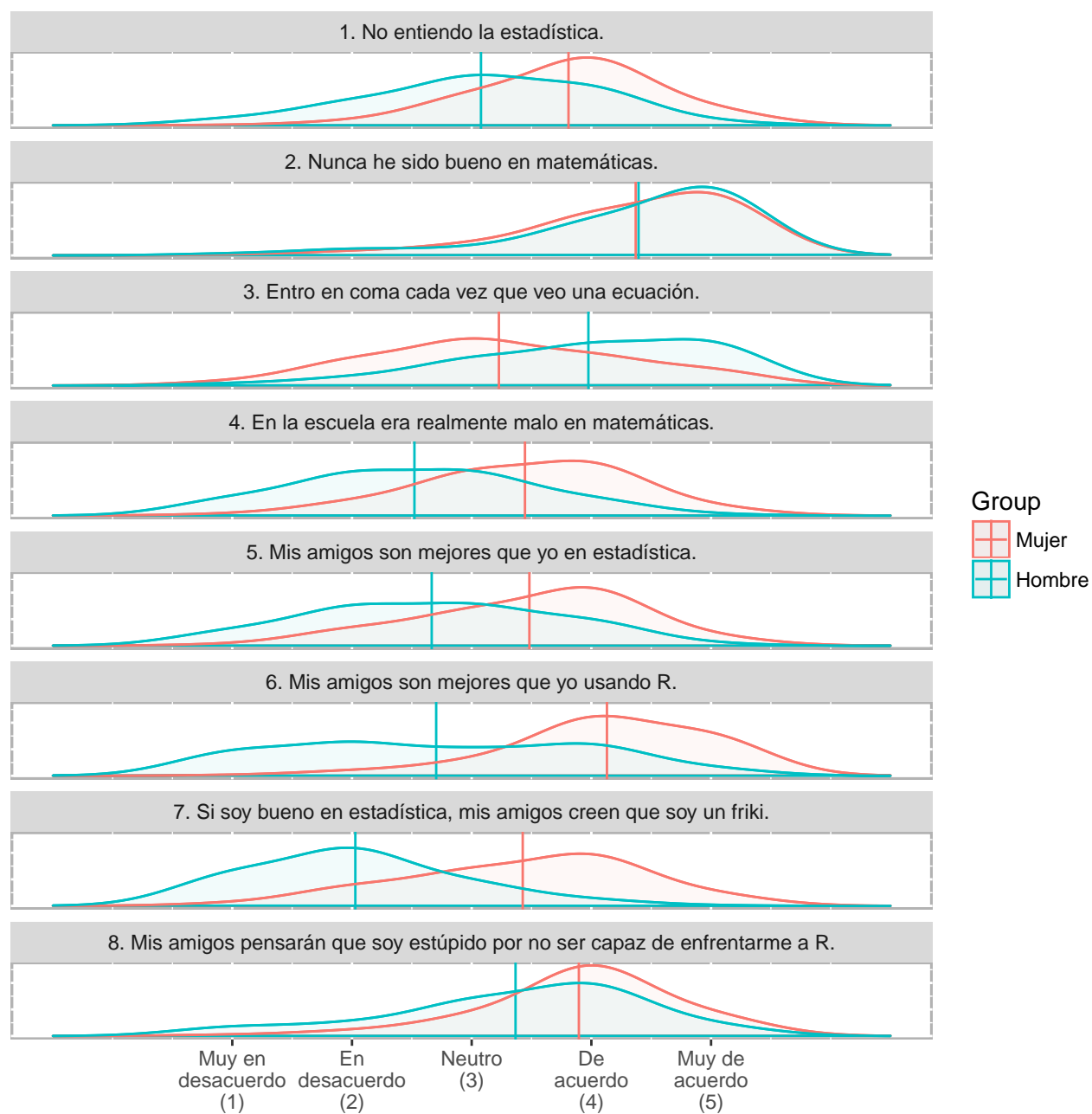



Figura 4: Gráfico de densidad según `sexo` de los ítems 1 a 8.

Referencias y bibliografía

Bryer, Jason, and Kimberly Speerscheider. 2015. *Likert: Functions to Analyze and Visualize Likert Type Items*. <http://CRAN.R-project.org/package=likert>.

Field, Andy, Jeremy Miles, and Zoe Field. 2012. *Discovering Statistics Using R*. 1st edition. Sage Publications Ltd.

Murdoch, Duncan. 2016. *Tables: Formula-Driven Table Generation*. <http://CRAN.R-project.org/package=>

tables.