

Escalamiento de Likert en R

Realizado por: Nahuel Antonio Salem García

Dirigido por: Teresa Rivas Moya

Introducción

La escala Likert (1932, citado en Rivas, 2014):

- es la más usada para medir actitudes
- gradúa el acuerdo con cada ítem en una escala subjetiva (habitualmente de 5 ó 7 puntos)
- requiere determinar el **valor de escala** de cada nivel de acuerdo para cada ítem

Los valores de escala se hallan a partir de las proporciones de sujetos que responden en cada categoría para ese ítem. Si la correlación entre los valores que se pretende usar y los valores de escala hallados es de 0.99 o mayor, la respuesta al ítem se puede considerar como **variable de intervalo** (Edwards, 1957; citado en Rivas, 2014). Originalmente este procedimiento se ha aplicado para ítems con 5 niveles de respuesta.

En este trabajo se explora:

- cómo visualizar datos de una escala Likert
- cómo obtener los valores de escala para una escala Likert.

Método

Los datos y la función expuestos a continuación están disponibles en el siguiente enlace: <https://github.com/NSalem/Likert/archive/master.zip>

Software

Para este trabajo se ha utilizado:

- El lenguaje y entorno de programación R versión 3.1.2 (R Core Team, 2014) y el paquete `likert` versión 1.2 (Bryer y Speerschneider, 2014) para el mismo.
- Una función de R que permite hallar los valores de escala para cada ítem de un conjunto de datos dado. Esta función se pone a prueba con un cuestionario de autoestima y otro de ansiedad ante las matemáticas.

Datos

Los datos que se utilizan son:

- **ae:** Las respuestas de 56 sujetos en un cuestionario de autoestima. 10 ítems que se puntúan en una escala de 4 puntos desde "Muy en desacuerdo" hasta "Muy de acuerdo". Datos extraídos de Rivas (2014).
- **mass:** Resultados de un cuestionario de ansiedad ante las matemáticas (*Math Anxiety Scale Survey*; Bai et al., 2009), administrado a 20 estudiantes en una clase de estadística. 6 ítems con una escala de 5 puntos. Obtenidos del paquete `likert`. Columna de sexo eliminada y nombres de ítems cambiados.

Visualización

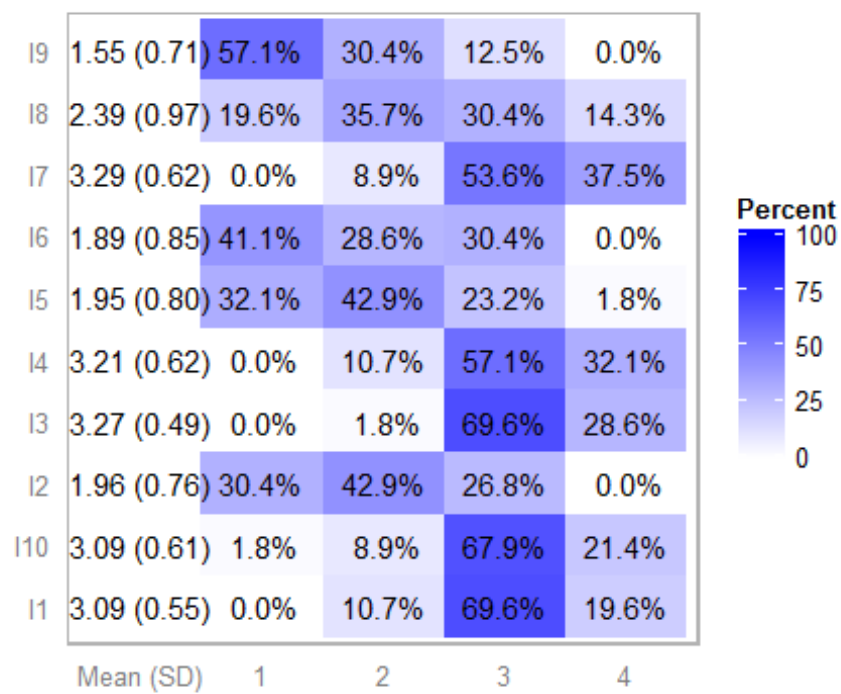
Una vez cargados los datos, se pueden visualizar con la ayuda del paquete `likert`:

- En caso de no tener el paquete instalado hay que usar el comando `install.packages("likert")` para hacerlo.
- Para cargar el paquete se usa `library("likert")`
- Usar la función `likert` para obtener los porcentajes. Si alguna proporción es 0, hay que especificar el número de niveles que queremos con el argumento `nlevels`
- Usar la función `plot` sobre los porcentajes.

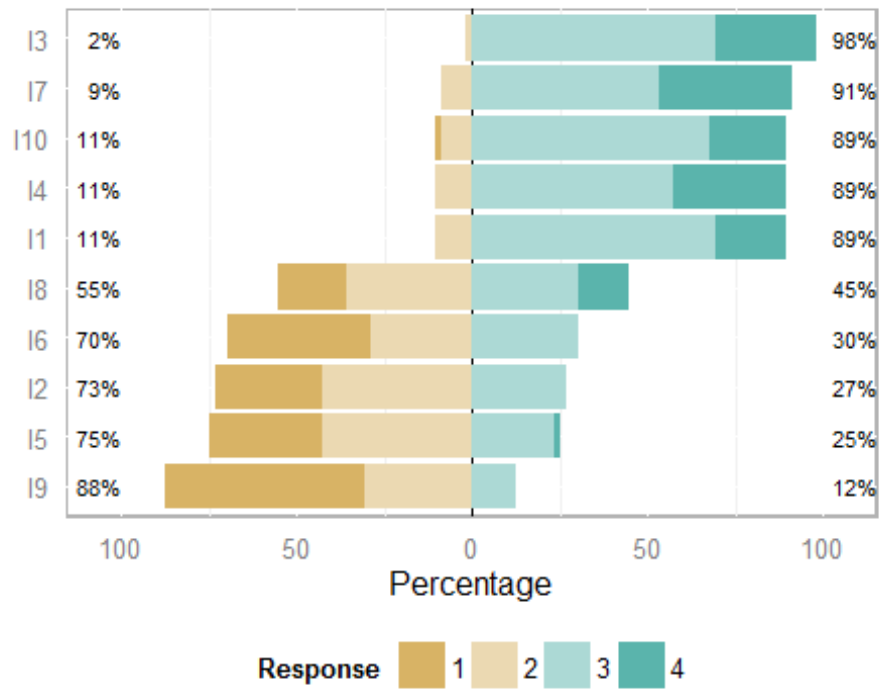
```
library("likert") #cargar paquete Likert
```

```
l_ae = likert(ae, nlevels = 4) #saca porcentajes, especificando que hay 4 niveles.
```

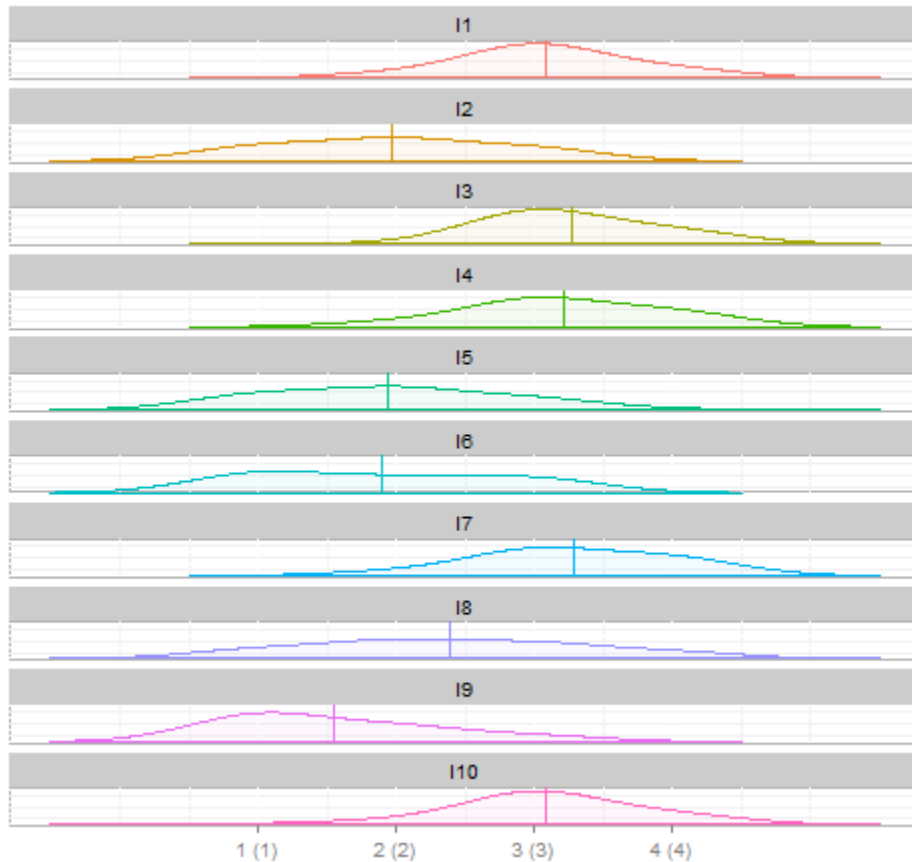
```
plot(l_ae, type = "heat") #tabla de porcentajes, media y dt para ítems
```



`plot(l_ae)` *#gráfico de barras, ver porcentajes*



```
plot(l_ae, type = "density") #grafico de densidad, ver respuesta al ítem como variable continua
```



En el primer y el segundo gráfico se pueden ver las frecuencias como porcentajes. Se ve que en algunos ítems la opción de respuesta 4 ("Totalmente de acuerdo") no representa ninguna proporción del total de respuestas al ítem. En otros ocurre lo mismo con la opción 1 ("Totalmente en desacuerdo").

El tercer gráfico permite ver la respuesta al ítem representada como una variable continua.

Valores de escala

A continuación se muestra cómo **hallar los valores de escala** mediante una función propia. Esta función, `v_escalas_likert`, toma una tabla de las respuestas dadas por varios sujetos (filas) a varios ítems (columnas) y devuelve los valores de escala para cada ítem y la correlación con los niveles que se pretende utilizar. La función admite dos formatos de datos para las respuestas a los ítems:

- **Numérico.** En este caso es necesario especificar entre qué valores está la escala de respuesta. Así, hay que escribir `v_escala_likert(datos, vmin, vmax)`, donde `datos` es el conjunto de datos, `vmin` es el número que representa el mínimo nivel de respuesta posible para los ítems y `vmax` es el número que representa el máximo nivel de respuesta posible para los ítems. Por ejemplo, en una escala de 5 puntos que va de 1 a 5, `vmin` debe ser 1 y `vmax` 5. Los valores por defecto para `vmin` y `vmax` son 1 y 5 respectivamente.
- **Factor ordenado** (las respuestas aparecen como “muy en desacuerdo”, “en desacuerdo”, etc y los niveles están ordenados. Ver ayuda de R escribiendo en consola `?factor`). `vmin` y `vmax` marcan el valor mínimo y máximo de la escala a numérica a la que se quieren traspasar, teniéndose que cumplir que `vmax - vmin` sea igual al número de categorías de respuesta. Es decir, si hay 5 categorías, los valores pueden ser tanto `vmin = 0, vmax = 4` como `vmin = 1, vmax = 5`.

```
v_escala_likert(ae, 1, 4)

##      1 2 3 4 correlaciones
## I1   NA 1 2 4             NA
## I2    1 2 3 NA             NA
## I3   NA 1 3 4             NA
## I4   NA 1 2 4             NA
## I5    1 2 3 4      1.0000000
## I6    1 2 3 NA             NA
## I7   NA 1 2 4             NA
## I8    1 2 3 4      1.0000000
## I9    1 2 3 NA             NA
## I10   1 2 3 5      0.9827076
```

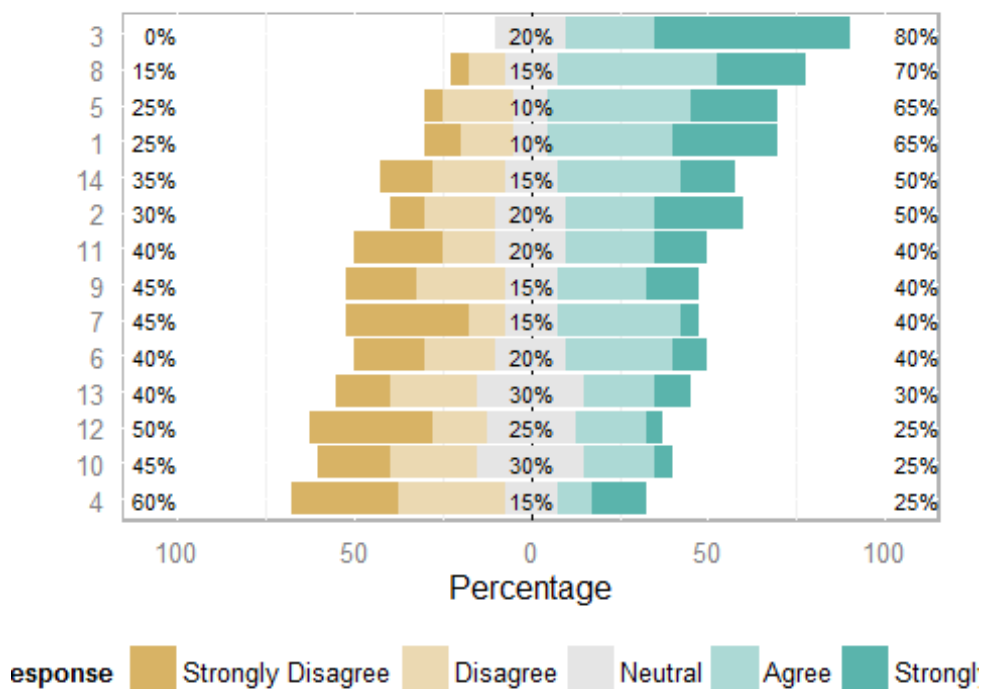
Así, se puede ver cómo en este cuestionario los valores de escala de diferentes ítems varían. Para muchos no se puede hallar el valor de escala correspondiente para algún nivel porque la proporción de respuestas en ese nivel es 0. Es Las únicas correlaciones mayores o iguales que .99 son las encontradas para los ítems 5 y 8.

Veamos ahora otro ejemplo con el cuestionario de ansiedad ante las matemáticas. Visualizar las proporciones en primer lugar puede ser útil. Habiendo cargado el paquete `likert` podemos introducir los siguientes comandos:

```
data(mass) # cargar los datos
mass = mass[-1] #quitar la columna de sexo
names(mass) = c(1:14) #OPCIONAL. Quitar texto de los ítems

#Nota: el archivo mass.RData contiene los datos ya transformados de
esta manera.

plot(likert(mass)) #hacer gráfico de barras
```



El ítem 3 no tiene ninguna proporción de respuestas en desacuerdo. Por tanto, no se podrán calcular los valores de escala para los dos primeros niveles de respuesta (“Fuertemente en desacuerdo” y “en desacuerdo”).

```
v_escala_likert(mass, 1, 5)
```

```
##      1  2 3 4 5 correlaciones
## 1    1  2 2 3 4      0.9707253
## 2    1  2 2 3 4      0.9707253
## 3    NA NA 1 2 3          NA
## 4    1  2 2 3 3      0.9449112
## 5    1  2 2 3 4      0.9707253
## 6    1  2 2 3 4      0.9707253
## 7    1  2 2 3 4      0.9707253
## 8    1  2 2 3 4      0.9707253
## 9    1  2 2 3 4      0.9707253
## 10   1  2 3 3 4      0.9707253
## 11   1  2 2 3 4      0.9707253
## 12   1  2 2 3 4      0.9707253
## 13   1  2 3 3 4      0.9707253
## 14   1  2 2 3 4      0.9707253
```

En este caso ninguna correlación llega a ser de 0.99 o más, por tanto no se puede considerar que la respuesta a ninguno de estos ítems sea una variable continua.

Al añadir `mostrar_todo=TRUE` o simplemente `TRUE` en la función `v_escala_likert` la función devuelve todos los pasos intermedios hasta llegar a los valores de escala:

```
v_escala_likert(ae,1, 4, TRUE)
```

```
## $`frecuencias relativas`
##           1           2           3           4
## 1  0.00000000 0.10714286 0.6964286 0.19642857
## 2  0.30357143 0.42857143 0.2678571 0.00000000
## 3  0.00000000 0.01785714 0.6964286 0.28571429
## 4  0.00000000 0.10714286 0.5714286 0.32142857
## 5  0.32142857 0.42857143 0.2321429 0.01785714
## 6  0.41071429 0.28571429 0.3035714 0.00000000
## 7  0.00000000 0.08928571 0.5357143 0.37500000
## 8  0.19642857 0.35714286 0.3035714 0.14285714
## 9  0.57142857 0.30357143 0.1250000 0.00000000
## 10 0.01785714 0.08928571 0.6785714 0.21428571
##
## $`frecuencias relativas acumuladas`
##           1           2           3 4
## 1  0.00000000 0.10714286 0.8035714 1
## 2  0.30357143 0.73214286 1.0000000 1
## 3  0.00000000 0.01785714 0.7142857 1
## 4  0.00000000 0.10714286 0.6785714 1
## 5  0.32142857 0.75000000 0.9821429 1
## 6  0.41071429 0.69642857 1.0000000 1
## 7  0.00000000 0.08928571 0.6250000 1
## 8  0.19642857 0.55357143 0.8571429 1
## 9  0.57142857 0.87500000 1.0000000 1
## 10 0.01785714 0.10714286 0.7857143 1
##
## $`media p acumulada`
##           1           2           3           4
## 1  0.00000000 0.053571429 0.4553571 0.9017857
## 2  0.151785714 0.517857143 0.8660714 1.0000000
## 3  0.000000000 0.008928571 0.3660714 0.8571429
## 4  0.000000000 0.053571429 0.3928571 0.8392857
## 5  0.160714286 0.535714286 0.8660714 0.9910714
## 6  0.205357143 0.553571429 0.8482143 1.0000000
## 7  0.000000000 0.044642857 0.3571429 0.8125000
## 8  0.098214286 0.375000000 0.7053571 0.9285714
## 9  0.285714286 0.723214286 0.9375000 1.0000000
## 10 0.008928571 0.062500000 0.4464286 0.8928571
##
## $`media p acumulada normalizada`
##           1           2           3           4
## 1          -Inf -1.61116916 -0.1121376 1.2917938
## 2 -1.0288048  0.04477618  1.1080108          Inf
## 3          -Inf -2.36856706 -0.3422764 1.0675705
```



```
## 4      -Inf -1.61116916 -0.2718800 0.9915265
## 5 -0.9915265 0.08964235 1.1080108 2.3685671
## 6 -0.8226373 0.13468979 1.0288048      Inf
## 7      -Inf -1.69917766 -0.3661064 0.8871466
## 8 -1.2917938 -0.31863936 0.5398714 1.4652338
## 9 -0.5659488 0.59241694 1.5341205      Inf
## 10 -2.3685671 -1.53412054 -0.1346898 1.2418668
##
## $`valores de escala`
##      1 2 3 4 correlaciones
## I1   NA 1 2 4              NA
## I2    1 2 3 NA              NA
## I3   NA 1 3 4              NA
## I4   NA 1 2 4              NA
## I5    1 2 3 4              1.0000000
## I6    1 2 3 NA              NA
## I7   NA 1 2 4              NA
## I8    1 2 3 4              1.0000000
## I9    1 2 3 NA              NA
## I10   1 2 3 5              0.9827076
```

Conclusiones

Con las figuras y los valores obtenidos podemos ver las diferencias en la distribución de las respuestas entre ítems. De la misma forma, podemos ver qué ítems pueden considerarse como variable de intervalo en cuanto a su respuesta.

En los ejemplos usados:

- Las respuestas a los **ítems 5 y 8** del cuestionario de autoestima se pueden considerar **variables continuas**. El resto de ítems no.
- Ningún ítem del cuestionario de ansiedad a las matemáticas se puede medir como una variable continua
- No se puede considerar la **puntuación total** en ninguno de los cuestionarios como variable continua.

Problemas encontrados:

- Niveles con proporciones de respuesta de 0
- Poca información encontrada sobre determinar valores de escala

Referencias

Bai, H., Wang, L., Pan, W., y Frey, M. (2009). Measuring mathematics anxiety: Psychometric analysis of a bidimensional affective scale. *Journal of Instructional Psychology*, 36 (3), 185- 193. <http://www.thefreelibrary.com/Measuring%20mathematics%20anxiety:%20psychometric%20analysis%20of%20a...-a0211235540>

Bryer, J. y Speersneider, K. (2014). likert: Functions to analyze and visualize likert type items. R package version 1.2. <http://CRAN.R-project.org/package=likert>

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rivas, T. Escalamiento Centrado en el Sujeto: Likert - Teoría. <http://psicologia.cv.uma.es/mod/resource/view.php?id=87356> (Técnicas de Investigación en Psicología Clínica y de la Salud. Campus Virtual UMA. 14-01-2014)