



CLASE: Software para estimar modelos de escalamiento clásico

Javier Revuelta¹, Vicente Ponsoda y Julio Olea
e-mail: javier.revuelta@uam.es
¹ Departamento de Psicología Social y Metodología
Universidad Autónoma de Madrid

RESUMEN

Este artículo presenta el programa *CLASE*, que permite estimar los modelos de escalamiento clásico para la tarea de comparaciones binarias (Dunn-Rankin y Thurstone), clasificación de objetos (intervalos aparentemente iguales e intervalos sucesivos) y la teoría de detección de señales. El programa está pensado como una herramienta didáctica para ser utilizado en el aula o de forma individual por los estudiantes. Por tanto, proporciona los resultados de estas técnicas según se describen en referencias clásicas, sin incorporar los últimos desarrollos de las mismas.

Palabras clave: escalamiento clásico, teoría de la detección de señales, software psicométrico.

ABSTRACT

This article presents the software *CLASE*, which estimates the classical scaling models for the task of paired comparisons (Dunn-Rankin and Thurstone), rating of objects (equal appearing and successive intervals) and the signal detection theory. The program is conceived as a didactic tool to be applied in classroom or by individual students, and for conducting small-scale research. Thus, it provides the results of the techniques as can be found in classical textbooks instead of incorporating the latest developments of these techniques.

Keywords: classical scaling, signal detection theory, psychometric software.

Nota

Esta investigación ha sido subvencionada por el proyecto de la DGICYT: PB 97-0049.



1.- Introducción

El programa *CLASE* (classical scaling program) permite estimar algunos de los modelos de escalamiento clásico aplicados en psicología, según se describen en Ponsoda (1986). En concreto, los modelos de escalamiento unidimensional y teoría de la detección de señales. En este artículo únicamente se describe el funcionamiento del programa. Los fundamentos de estas técnicas pueden consultarse en las referencias citadas.

La finalidad principal del programa es servir de apoyo a la docencia, por lo que no incorpora muchos desarrollos recientes de estas técnicas (Bock y Jones, 1968). No obstante, esto no impide que también pueda resultar útil para analizar datos en investigaciones reales.

Con el objetivo de facilitar la labor del usuario, *CLASE* permite leer dos tipos de archivos de datos: un archivo con los patrones de respuesta de cada sujeto o con las frecuencias de respuesta. De esta manera, pueden prepararse los datos de la forma que resulte más cómoda, según sea su formato original.

La forma concreta de preparar el archivo de datos depende de la técnica que vaya a aplicarse por lo que los detalles se comentan en los capítulos referidos a las técnicas. Los archivos de datos deben llevar la extensión '.dat' y *CLASE* crea un archivo de resultados con el mismo nombre que el de datos y la extensión '.out'. Tanto los archivos de datos como los de resultados tienen formato texto.

El programa *CLASE* se utiliza a partir de la línea de comandos de MS-DOS. Supongamos que en el directorio 'C:\Experimentos' se dispone de un archivo de datos llamado 'Comparaciones.dat' junto con el programa *CLASE*. Para realizar el análisis es necesario abrir una ventana de MS-DOS y ejecutar la orden:

```
C:\Experimentos> Clase 1 Comparaciones
```

El número 1 indica el análisis que debe realizarse, en concreto, los valores posibles son:

- 1: Escalamiento clásico, tarea de comparaciones binarias.
- 2: Escalamiento clásico, tarea de clasificación.
- 3: Teoría de la detección de señales.

A partir de esta orden, *CLASE* realiza el análisis y crea el archivo de resultados 'Comparaciones.out'.

2.- Escalamiento unidimensional

2.1.- Tipos de análisis

CLASE permite analizar los datos de dos tareas experimentales (Edwards y Thurstone, 1952; Edwards, 1957):



- Tarea de *comparaciones binarias*. Análisis mediante los métodos de Dunn-Rankin y Thurstone.
- Tarea de *clasificación de los estímulos*. Análisis mediante los métodos de intervalos aparentemente iguales e intervalos sucesivos.

2.2.- Tarea de comparaciones binarias

En una tarea de comparaciones binarias se le presentan a cada sujeto cada una de las $N(N-1)/2$ binas que pueden formarse con un conjunto de N objetos. Para cada par de objetos el sujeto debe indicar cual es el que prefiere, o en general, cual posee un mayor valor en alguna característica de interés.

2.2.1.- Archivos de datos

Existen dos tipos de archivos de datos: archivos con patrones de respuesta y archivos con frecuencias de respuesta. El formato en que se presentan los datos debe indicarse dentro del archivo de datos.

2.2.1.1.- Archivo de patrones de respuesta

Contiene las respuestas tal cual las dan los sujetos. Supongamos que se realiza un experimento con varios sujetos y un conjunto de 4 objetos. Entonces a cada sujeto se le presentan 6 pares de objetos. El archivo de datos tiene el aspecto:

```
1 4
1 1 1 1 1 1
1 1 1 1 1 2
1 1 1 1 2 1
1 1 1 1 2 2
1 1 1 2 1 1
```

El primer valor (1) indica que se trata de un archivo de patrones de respuesta, el segundo valor (4) indica que hay cuatro objetos en el conjunto.

A continuación se muestran los resultados para los cinco primeros sujetos. Cada sujeto aparece en una fila y cada par en una columna. La primera columna corresponde al par (1, 2), la segunda al par (1, 3), la tercera al par (1, 4), la cuarta al par (2, 3), la quinta al par (2, 4) y la sexta al par (3, 4). El primer sujeto ha escogido el primer objeto de cada par en las seis presentaciones. El segundo sujeto ha escogido el primer objeto en todos los pares menos el último, en el cual ha escogido el segundo, etc.

2.2.1.2.- Archivo de frecuencias de respuesta

Contiene la matriz de comparaciones binarias. Supongamos que se realiza el experimento con 200 sujetos y 4 objetos. Los resultados podrían ser (Ponsoda, 1986, pág. 37):



2	4			
0	20	180	190	
180	0	160	170	
20	40	0	150	
10	30	50	0	

El primer valor (2) indica que se trata de un archivo de frecuencias de respuesta, el valor 4 que hay cuatro objetos. A continuación aparecen las frecuencias con que se ha escogido cada objeto de cada par. Los ceros en la diagonal corresponden a los pares (1, 1) a (4, 4), que son pares que no se aplican nunca. El valor 20 indica que en el par (1, 2) ha habido 20 sujetos que han preferido el objeto 1 sobre el 2, etc.

2.2.2.- Archivo de resultados

El archivo de resultados de *CLASE* es igual para los dos tipos de archivos de datos. En el ejemplo del archivo de frecuencias de respuesta su aspecto es:

Binary comparisons:											
Frequencies				Proportions				Normal deviates			
0	20	180	190	0.000	0.100	0.900	0.950	0.000	-1.282	1.282	1.645
180	0	160	170	0.900	0.000	0.800	0.850	1.282	0.000	0.842	1.036
20	40	0	150	0.100	0.200	0.000	0.750	-1.282	-0.842	0.000	0.674
10	30	50	0	0.050	0.150	0.250	0.000	-1.645	-1.036	-0.674	0.000

Scale values		
Dunn Rankin	Thurstone	
65.0000	1.2502	
85.0000	1.6288	
35.0000	0.4768	
15.0000	0.0000	

En primer lugar aparece la matriz de comparaciones binarias. Después la matriz de comparaciones convertida en proporciones y la matriz de puntuaciones normales. A continuación se muestran los valores de cada objeto en la escala psicológica para el método de Dunn Rankin y el de Thurstone. Para el método de Thurstone se toma como referencia el último objeto, por lo que su valor de escala es siempre 0.

2.3.- Tarea de clasificación de los estímulos

En la tarea de clasificación de estímulos se le presentan al sujeto los N objetos uno a uno para que los valore en una escala con C categorías de respuesta.

2.3.1.- Archivos de datos

Los datos pueden introducirse en forma de patrones o de frecuencias, aunque la forma exacta de hacerlo difiere con respecto a la tarea de comparaciones binarias.



2.3.1.1.- Archivo de patrones de respuesta

Supongamos que a cada sujeto se le presentan 4 estímulo y debe valorarlos en una escala de 5 puntos. El archivo de patrones contiene la matriz de respuestas, donde los sujetos están en las filas y los objetos en las columnas. Un posible archivo sería:

```
1 4
1 2 3 5
5 1 4 4
1 4 2 5
4 3 2 5
2 3 4 4
```

El primer valor 1 indica que se trata de un archivo de patrones de respuesta y el 4 que en el experimento hay cuatro objetos. A continuación aparecen las valoraciones de los cinco primeros sujetos, una en cada fila.

No es necesario decirle a *CLASE* cual es el número de valores de la escala. El programa entiende que esta escala está comprendida entre 1 y el máximo valor que aparezca en los datos.

2.3.1.2. -Archivo de frecuencias de respuesta

El archivo de frecuencias contiene el número de sujetos que han concedido cada valoración a cada objeto. En este caso es necesario decirle a *CLASE* cuantos niveles tiene la escala pero no cuantos objetos hay. Un ejemplo es el siguiente (Ponsoda, 1986, pág 47):

```
2 7
0 0 10 20 50 10 10
10 10 30 30 10 10 0
0 10 10 30 30 10 10
0 10 30 40 20 0 0
0 10 40 30 20 0 0
```

El valor 2 indica que se trata de un archivo de frecuencias y el 7 que la escala consta de 7 niveles. A continuación aparecen los resultados para cinco objetos, uno en cada fila. *CLASE* entiende que hay tantos objetos como filas aparezcan en este archivo, después de la cabecera de la primera fila. Los datos de cada fila indican cuantos sujetos han clasificado cada objeto en cada categoría. La categoría 1 corresponde a la columna izquierda y la 7 a la derecha

2.3.2.- Archivo de resultados

El archivo de resultados tiene un formato común para ambos tipos de datos. En primer lugar muestra los resultados para el método de intervalos aparentemente iguales. En el ejemplo del archivo con frecuencias de respuesta:



Equal Appearing Intervals:

Object value	Frec. 1	Frec. 2	Frec. 3	Frec. 4	Frec. 5	Frec. 6	Frec. 7	Scale
1	0	0	10	20	50	10	10	4.900
2	10	10	30	30	10	10	0	3.500
3	0	10	10	30	30	10	10	4.500
4	0	10	30	40	20	0	0	3.750
5	0	10	40	30	20	0	0	3.500

Esta tabla muestra la frecuencia de las clasificaciones de cada objeto y, en la columna derecha, los valores de escala. A continuación aparecen los resultados para el método de intervalos sucesivos:

Successive Intervals:

Proportions and normal deviates

0.00	0.00	0.10	0.30	0.80	0.90	1.00	0.000	0.000	-1.282	-0.524	0.842	1.282
0.10	0.20	0.50	0.80	0.90	1.00	1.00	-1.282	-0.842	0.000	0.842	1.282	0.000
0.00	0.10	0.20	0.50	0.80	0.90	1.00	0.000	-1.282	-0.842	0.000	0.842	1.282
0.00	0.10	0.40	0.80	1.00	1.00	1.00	0.000	-1.282	-0.253	0.842	0.000	0.000
0.00	0.10	0.50	0.80	1.00	1.00	1.00	0.000	-1.282	0.000	0.842	0.000	0.000

Interval width

0.000	0.000	0.000	0.757	1.366	0.440	0.000
0.000	0.440	0.842	0.842	0.440	0.000	0.000
0.000	0.000	0.440	0.842	0.842	0.440	0.000
0.000	0.000	1.028	1.095	0.000	0.000	0.000
0.000	0.000	1.282	0.842	0.000	0.000	0.000

Mean width and scale

0.000	0.440	0.898	0.875	0.883	0.440	1.000
0.000	0.440	1.338	2.213	3.096	3.536	4.536

Scale value

2.563
1.338
2.213
1.562
1.338

CLASE muestra las proporciones acumuladas de las clasificaciones y estas proporciones convertidas en valores normales. A continuación aparece la anchura de cada intervalo para cada uno de los cinco objetos. Después la anchura media de los intervalos y la anchura media convertida en valores de escala. Por último, los valores de escala de los objetos.

3.- Teoría de Detección de Señales

En la Teoría de la Detección de Señales (TDS) se le presentan al sujeto varios estímulos de dos tipos diferentes (conocidos como señal o ruido) y su tarea consiste en clasificarlos en dos o más grupos distintos. Para cada sujeto se obtiene un valor d' (índice de sensibilidad) que indica el grado en que el sujeto es capaz de percibir la diferencia entre los



dos tipos de estímulos y uno o más valores β (criterio) que indican la tendencia del sujeto para clasificar los estímulos en cada una de las posibles categorías (Green y Sweets, 1966).

Los datos de la TDS deben contener la información del tipo de estímulo y de respuesta. Además, *CLASE* puede leer la matriz de consecuencias del experimento, la cual determina cual es la consecuencia de cada una de las posibles clasificaciones y se utiliza para determinar el valor óptimo (el que maximiza las consecuencias) de β .

3.1.- Análisis con patrones de respuesta

Los datos contienen las respuestas individuales de cada sujeto a cada estímulo.

3.1.1.- Datos dicotómicos

En este caso el sujeto sólo puede clasificar los estímulos en dos categorías. Un ejemplo sería el siguiente:

```

1 2 24
1 1
1 2
2 1
2 2
1 1

```

El valor 1 indica que se trata de patrones de respuesta, el 2 que hay dos categorías de respuesta y el 24 que a cada sujeto se le han presentado 24 estímulos. Después aparecen los resultados para los cinco primeros estímulos del primer sujeto. La primera columna indica el tipo de estímulo y la segunda el tipo de respuesta.

Dado que se han presentado 24 estímulos, el número de filas que aparece después de la cabecera inicial debe ser un múltiplo de 24. Cada bloque de 24 filas corresponde a un sujeto.

Supongamos que se realiza un experimento en el que el número de estímulos es diferente para cada sujeto. Entonces habría que preparar los archivos de datos por separado y analizarlos mediante *CLASE* de forma independiente. El programa entiende que a todos los sujetos que aparecen en un mismo archivo de datos se les ha aplicado el mismo número de estímulos.

El archivo de resultados tiene el aspecto:

Signal Detection Theory:

```

Subject      1
  10   4   0.714 0.286  -0.566
   4   6   0.400 0.600   0.253

PA          :   0.6571
d prime     :   0.8193
Beta 1     :   0.8798
Number Correct Optimum Beta :   0.7143

```



```
Subject      2
  7  6  0.538 0.462 -0.097
  6  5  0.545 0.455 -0.114

PA          : 0.4965
d prime    : -0.0176
Beta 1     : 1.0019
Number Correct Optimum Beta : 0.8462
```

Los resultados se muestran sujeto por sujeto. En primer lugar aparece la tabla de contingencia de sus respuestas. Después las proporciones correspondientes a dicha tabla. En la columna derecha están los valores normales de la tabla de contingencia.

A continuación están los resultados. En primer lugar PA, que es un índice no paramétrico de la sensibilidad. Después d' y β . Por último se muestra el valor óptimo de β para el que se maximiza el número de respuestas correctas.

3.1.2.- Datos dicotómicos y una matriz de consecuencias

En caso de que desee obtenerse el valor óptimo de β para una determinada matriz de consecuencias, habría que indicarlo en el archivo de datos del siguiente modo:

```
3 2 24
 5 -5
-2  2
1  1
1  2
2  1
2  2
1  1
```

El valor 3 indica que se trata de un archivo de patrones de respuesta más una matriz de consecuencias. Los valores 2 y 24 tienen el sentido ya indicado. A continuación aparece la matriz de consecuencias (filas segunda y tercera del archivo de datos) y después los patrones de respuesta para los sujetos.

El archivo de resultados tiene el mismo aspecto que en el caso anterior, aunque se ha añadido la información referida a esta matriz de consecuencias:

Signal Detection Theory:

```
Subject      1
 10  4  0.714 0.286 -0.566
   4  6  0.400 0.600  0.253

PA          : 0.6571
d prime    : 0.8193
Beta 1     : 0.8798
```



Consequences:

```
5.0 -5.0
-2.0 2.0
Consequences Optimum Beta : 0.2857
Number Correct Optimum Beta : 0.7143
```

Puede verse que *CLASE* ha añadido la matriz de consecuencias y el valor óptimo de β para dicha matriz.

3.1.3.- Datos politómicos

Se obtienen cuando el sujeto puede clasificar cada estímulo en más de dos categorías. Un ejemplo sería el siguiente:

```
1 5 24
1 5
1 2
2 1
2 2
1 1
```

El valor 1 indica que se trata de patrones de respuesta, 5 es el número de categorías y 24 el número de estímulos por sujeto. A continuación pueden verse los resultados para los cinco primeros estímulos presentados al primer sujeto.

El archivo de resultados tiene el aspecto:

Signal Detection Theory:

```
Subject      1
3 4 2 1 4   0.214 0.286 0.143 0.071 0.286 0.792 0.000 -0.366 -0.566
3 5 0 1 1   0.300 0.500 0.000 0.100 0.100 0.524 -0.842 -0.842 -1.282

PA           : 0.3643
d prime      : -0.5750
Beta 1       : 0.8387
Beta 2       : 1.4250
Beta 3       : 1.3326
Beta 4       : 1.9368
```

En primer lugar está la tabla de contingencia con las frecuencias de las clasificaciones en las cinco categorías. Después la tabla en proporciones y la tabla en puntuaciones normales. Los resultados incluyen los índices PA y d' . Además pueden verse los cuatro valores de β que corresponden a las cinco categorías en que se clasifican los estímulos.

En caso de utilizar datos politómicos, *CLASE* no proporciona los valores óptimos de β ni puede leer una matriz de consecuencias.



3.2.- Análisis con frecuencias de respuesta

Los datos contienen las frecuencias de las clasificaciones de cada sujeto. El aspecto del archivo de resultados es el mismo que en el caso anterior, por lo que no va a hacerse ningún comentario adicional sobre el mismo.

3.2.1.- Datos dicotómicos

Un ejemplo es el siguiente:

```
2 2
80 20 10 90
94 6 44 56
90 10 48 52
```

El primer valor 2 indica que se trata de un archivo de frecuencias de respuesta y el segundo que los datos son dicotómicos. A continuación aparecen los datos de tres sujetos, cada uno en una fila. Cada columna de la tabla indica la frecuencia con que el sujeto ha asociado cada estímulo a cada respuesta. La primera columna corresponde al estímulo 1 y la respuesta 1, la segunda al estímulo 1 y la respuesta 2, después está el estímulo 2 y la respuesta 1 y por último el estímulo 2 y la respuesta 2.

3.2.2.- Datos dicotómicos y una matriz de consecuencias

Añade al archivo anterior la existencia de una matriz de consecuencias. Un ejemplo es el siguiente:

```
4 2
10 -4
-9 7
80 20 10 90
94 6 44 56
90 10 48 52
```

El valor 4 indica que es un archivo de frecuencias con una matriz de consecuencias. La matriz aparece en la segunda y tercera filas.

3.2.3.- Datos politómicos

Cada objeto puede clasificarse en más de dos categorías. Los datos tienen el aspecto:

```
2 5
60 20 6 8 6 4 8 12 20 56
10 24 36 28 14 39 31 23 29 41
```

El valor 2 indica que se trata de frecuencias de respuesta y el 5 es el número de posibles respuestas. A continuación aparecen los datos para dos sujetos, uno en cada fila. Las cinco columnas de la izquierda se refieren al estímulo 1 y las respuestas 1 a 5. Las cinco columnas de la derecha se refieren al estímulo 2 y las respuestas 1 a 5.



4.- Referencias

- Bock, R. D. y Jones, L. V. (1968). *The measurement and prediction of judgement and choice*. San Francisco: Holden-Day.
- Edwards, A. L. (1957). *Techniques for attitude scale construction*. New York: Appleton-Century-Crofts.
- Edwards, A. L. y Thurstone, L. L. (1952). An internal consistency check for scale values determined by the method of successive intervals. *Psychometrika*, 17 (2) 169-180.
- Green, D. M. y Sweets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Krieger.
- Ponsoda, V. (1986). *Iniciación a la Psicología Matemática*. Publicaciones de la Universidad Autónoma de Madrid: Madrid.