



## Análisis de datos para la toma de decisiones

---

Profesora: Katherine Barquero Mejías



## Clase 3: Reducción de la dimensionalidad

# Reducción de la dimensionalidad

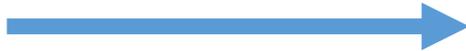
- En la práctica es común tener mucha información a partir de la cual se requieren simplificar los datos.
- Usualmente tenemos varios factores a partir de los cuales medimos una o varias cosas.
- Esto es común en muchos tipos de datos y en particular en los que generamos a partir de cuestionarios. Cuando trabajamos con **escalas**
- Aquí vamos a tener un

# Escalas

- Serie de valores o números ordenados que se utilizan para medir o cuantificar una variable.
- En general tenemos la posibilidad de encontrar escalas ordinales y escalas nominales.

## Escala nominal

- No tiene un orden específico
- Usualmente la variable se construye sumando los ítems que la componen



V18. ¿Tienen en esta vivienda...	
	¿Cuántos?
	Sí      ↓      No
V18a.	...teléfono celular? ..... <input type="radio"/> 1 <input type="checkbox"/> <input type="radio"/> 2
V18b.	...teléfono residencial (telefonía fija)? ..... <input type="radio"/> 3 <input type="checkbox"/> <input type="radio"/> 4
V18c.	...refrigeradora? ..... <input type="radio"/> 5 <input type="checkbox"/> <input type="radio"/> 6
V18d.	...sistema de agua caliente para toda la casa? ..... <input type="radio"/> 7 <input type="checkbox"/> <input type="radio"/> 8
V18e.	...tanque para almacenar agua? ..... <input type="radio"/> 1 <input type="checkbox"/> <input type="radio"/> 2
V18f.	...computadora portátil? ..... <input type="radio"/> 3 <input type="checkbox"/> <input type="radio"/> 4
V18g.	...computadora de escritorio? .. <input type="radio"/> 5 <input type="checkbox"/> <input type="radio"/> 6
V18q.	...tableta? ..... <input type="radio"/> 1 <input type="checkbox"/> <input type="radio"/> 2
V18h.	...fax (no de computadora)? ... <input type="radio"/> 7 <input type="checkbox"/> <input type="radio"/> 8
V18i.	...radio o equipo de sonido? ... <input type="radio"/> 1 <input type="checkbox"/> <input type="radio"/> 2
V18j.	...carro (no de trabajo)? ..... <input type="radio"/> 3 <input type="checkbox"/> <input type="radio"/> 4
V18k.	...moto (no de trabajo)? ..... <input type="radio"/> 5 <input type="checkbox"/> <input type="radio"/> 6
V18o.	...servicio de televisión paga por cable? ..... <input type="radio"/> 3 <input type="checkbox"/> <input type="radio"/> 4
V18o1.	¿De que empresa?

# Escalas

## Escala ordinal

- Tiene un orden específico
- Usualmente la variable se construye **promediando** los ítems que la componen
- Escala de Likert (1-5) o (1 a 7) para medir actitudes u opiniones de las personas

Ahora, vamos a usar una escalera en donde el número 1 representa "muy en desacuerdo" y el número 7 representa "muy de acuerdo". Un número entre el 1 y el 7, representa un puntaje intermedio.

1	2	3	4	5	6	7	888888	988888
Muy en desacuerdo			Muy de acuerdo			No sabe [NO LEER]	No responde [NO LEER]	

[Anotar un número 1-7, 888888 = No sabe, 988888= No responde]

Le voy a leer algunas frases. Por favor dígame hasta qué punto está de acuerdo o en desacuerdo con ellas.

PPL1. Los políticos siempre deberían escuchar con atención a los problemas del pueblo.	
PPL2. Los políticos no necesitan pasar tiempo con la gente común para hacer un buen trabajo.	
PPL3. La voluntad del pueblo debería ser el asunto más importante en la política de este país.	
ANT1. El gobierno está dirigido principalmente por algunos grandes intereses que sólo velan por sí mismos.	
ANT2. Los funcionarios del gobierno usan su poder para tratar de mejorar la vida de la gente.	
ANT3. Muchas de las personas que dirigen el gobierno son deshonestas.	
MAN1. Usted puede saber si una persona es buena o mala si conoce sus posiciones políticas.	
MAN2. Las personas que tienen ideas políticas distintas a las mías no son malas.	
MAN3. Las personas con las que no estoy de acuerdo políticamente sólo están mal informadas.	

**VB50.** Algunos dicen que en general, los hombres son mejores líderes políticos que las mujeres. ¿Está usted muy de acuerdo, de acuerdo, en desacuerdo, o muy en desacuerdo?

(1) Muy de acuerdo

(2) De acuerdo

(3) En desacuerdo

(4) Muy en desacuerdo

(888888) No sabe [NO LEER]

(988888) No responde [NO LEER]

# Criterios de las escalas

- Las escalas deben cumplir con dos criterios importantes antes de ser resumidas en una sola variable: ser confiables y unidimensionales
- Confiables: se estima un indicador estadístico llamado Alpha de Cronbach's que mide la consistencia interna dentro de la Escala. Toma valores entre 0 y 1.

Valores de interpretación Alpha Chronbach's	
Valor	Confiabilidad
0.5 o menos	Confiabilidad nula
0.5 a 0.6	Confiabilidad baja
0.6 a 0.7	Confiabilidad aceptable
0.7 a 0.8	Confiabilidad alta
Más de 0.8	Confiabilidad muy alta

- Unidimensionalidad: Se realiza análisis factorial exploratorio y se verifica que los ítems carguen en un solo factor. Lo que permite luego crear una variable con varios ítems de la escala

# Veamos el ejemplo de la Escala PER en Stata

Ahora, voy a leerle una serie de rasgos de personalidad que podrían aplicarse o no aplicarse a usted. Por favor use la misma escalera del 1 al 7 para indicar en qué medida está de acuerdo o en desacuerdo en que estas frases se aplican a su persona. Debe calificar en qué medida se aplican a usted estos rasgos de personalidad, aun cuando alguna característica se aplique en mayor medida que otra.

1	2	3	4	5	6	7	888888	988888
Muy en desacuerdo				Muy de acuerdo			No sabe [NO LEER]	No responde [NO LEER]

**Usted se considera una persona que es:**

**PER1.** Sociable y activa

**PER2.** Una persona crítica y peleonera

**PER3.** Una persona confiable y disciplinada

**PER4.** Una persona ansiosa y fácil de molestarse

**PER5.** Una persona abierta a nuevas experiencias e intelectual

**PER6.** Una persona callada y tímida

**PER7.** Una persona generosa y cariñosa

**PER8.** Una persona desorganizada y descuidada

**PER9.** Una persona calmada y emocionalmente estable

**PER10.** Una persona poco creativa y con poca imaginación

Cuestionario  
LAPOP 2018

# Estimación del Alpha de Cronbach's

Local projections for IRFs  
Model selection for ARIMA and ARFIMA  
New spline functions  
Corrected and consistent AICs  
IV fractional probit model  
IV quantile regression

Data Editor enhancements  
Do-file Editor enhancements  
New reporting features  
Graph colors by variable  
All-new graph style

[and more](#)

Possible actions

[Install available updates](#) (or type `-update all-`)

Click to [edit automatic update checking](#) preferences

```
. import excel "C:\Users\Katherine Barquero\OneDrive\ICAP\CursoAnalisisdatos
> eIndicadores\Clase3_18082023\Presentaciones\DatosEscalaPER.xlsx", sheet("C
> ostaRica-LAPOP-2018") firstrow

. alpha per1 per2 per3 per4 per5 per6 per7 per8 per9 per10

Test scale = mean(unstandardized items)
Reversed items: per2 per4 per6 per8 per10

Average interitem covariance:      .296806
Number of items in the scale:      10
Scale reliability coefficient:      0.5262

.
```

Nombre	Etiqueta
Encuestado	Encuestado
per1	per1
per2	per2
per3	per3
per4	per4
per5	per5
per6	per6
per7	per7
per8	per8
per9	per9
per10	per10

Propiedades

- Variables
  - Nombre
  - Etiqueta
  - Typo
  - Formato
  - Etiqueta de valor
  - Notas
- Datos
  - Nombre de archivo
  - Etiqueta
  - Notas
  - Variables
  - Observaciones
  - Tamaño
  - Memoria
  - Ordenado por

Comando



# Estimación del Alpha de Cronbach's

Local projections for IRFs  
Model selection for ARIMA and ARFIMA  
New spline functions  
Corrected and consistent AICs  
IV fractional probit model  
IV quantile regression

Data Editor enhancements  
Do-file Editor enhancements  
New reporting features  
Graph colors by variable  
All-new graph style

[and more](#)

Possible actions

[Install available updates](#) (or type -update all-)

Click to [edit automatic update checking](#) preferences

```
. import excel "C:\Users\Katherine Barquero\OneDrive\ICAP\CursoAnalisisdatos  
> eIndicadores\Clase3_18082023\Presentaciones\DatosEscalaPER.xlsx", sheet("C  
> ostaRica-LAPOPOP-2018") firstrow  
  
. alpha per1 per2 per3 per4 per5 per6 per7 per8 per9 per10  
  
Test scale = mean(unstandardized items)  
Reversed items: per2 per4 per6 per8 per10  
  
Average interitem covariance: .296806  
Number of items in the scale: 10  
Scale reliability coefficient: 0.5262  
  
.
```

Nombre	Etiqueta
Encuestado	Encuestado
per1	per1
per2	per2
per3	per3
per4	per4
per5	per5
per6	per6
per7	per7
per8	per8
per9	per9
per10	per10

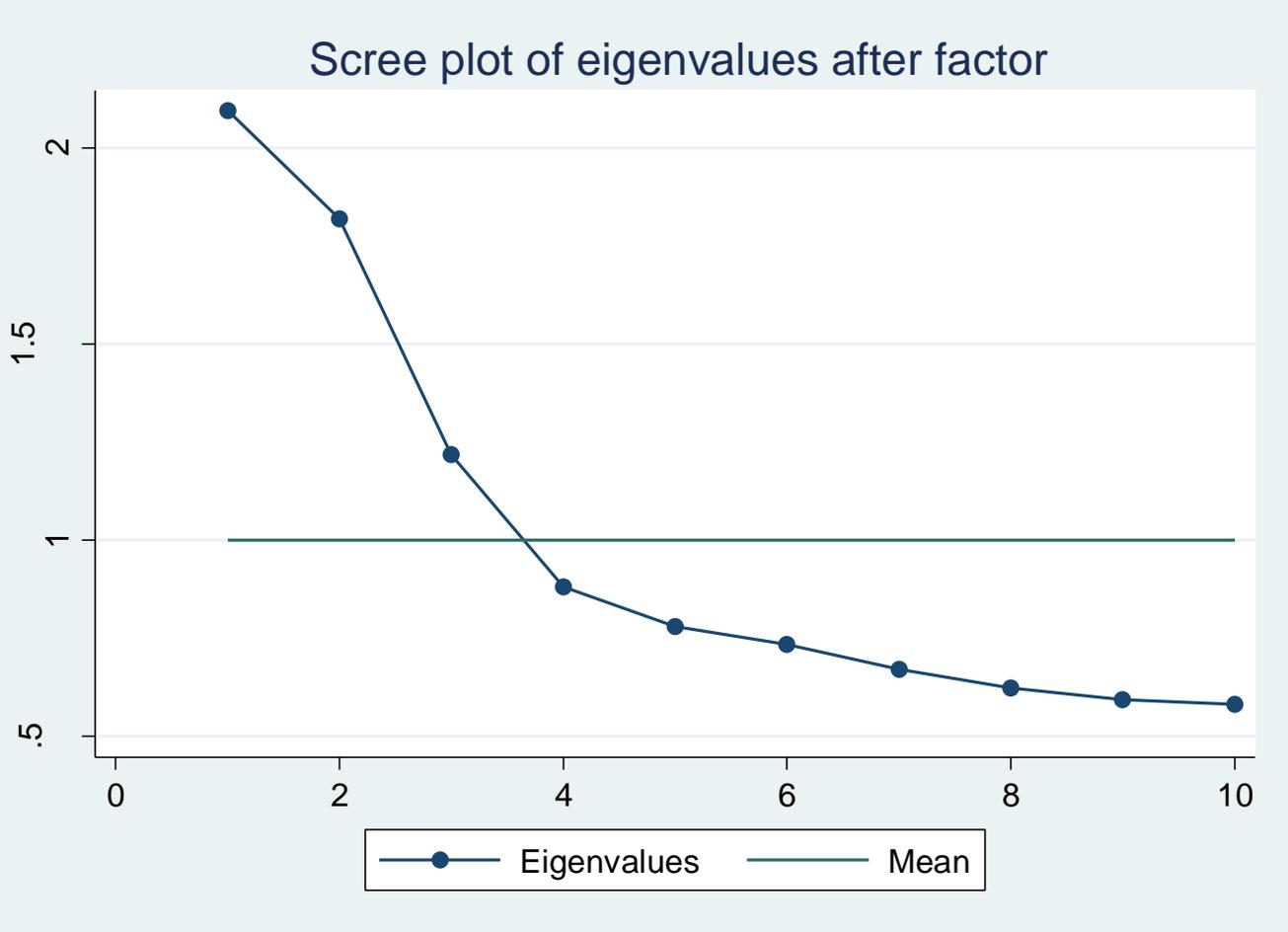
Propiedades

- Variables
  - Nombre
  - Etiqueta
  - Typo
  - Formato
  - Etiqueta de valor
  - Notas
- Datos
  - Nombre de archivo
  - Etiqueta
  - Notas
  - Variables
  - Observaciones
  - Tamaño
  - Memoria
  - Ordenado por

# Análisis factorial exploratorio

Primero se genera el Screeplot o gráfico de sedimentación: ¿Cuántos factores usar?

factor per1 per2 per3 per4 per5 per6 per7 per8 per9 per10 , pcf  
screeplot, mean



# Análisis factorial exploratorio

Se estiman las cargas factoriales. Primero corremos en un solo factor (una sola variable) para valorar las cargas factoriales y luego comparamos con el lo sugerido por el screeplot

factor per1 per2 per3 per4 per5 per6 per7 per8 per9 per10 , factor (1)  
rotate, oblique promax

```
LR test: independent vs. saturated:  ch12(45) = 1578.87  Prob>ch12 = 0.00  
> 00
```

Rotated factor loadings (pattern matrix) and unique variances

Variable	Factor1	Uniqueness
per1	0.5261	0.7233
per2	-0.0033	1.0000
per3	0.5140	0.7358
per4	-0.0265	0.9993
per5	0.4702	0.7789
per6	-0.0040	1.0000
per7	0.5466	0.7013
per8	-0.1907	0.9636
per9	0.4127	0.8297
per10	-0.1237	0.9847

Factor rotation matrix

	Factor1
Factor1	1.0000

The image shows a portion of the SPSS software interface. On the right side, there is a sidebar with several sections. At the top, there is a search bar labeled 'Filtrar variables'. Below it is a list of variables: 'per1', 'per2', 'per3', 'per4', 'per5', 'per6', 'per7', 'per8', 'per9', and 'per10'. Below the variable list is a section titled 'Propiedades' (Properties) with a lock icon and a plus sign. Under 'Propiedades', there are two expandable sections: 'Variables' and 'Datos'. The 'Variables' section is expanded and shows a list of properties for the selected variables: 'Nombre', 'Etiqueta', 'Tipo', 'Formato', 'Etiqueta d', and 'Notas'. The 'Datos' section is also expanded and shows a list of properties: 'Nombre c', 'Etiqueta', 'Notas', 'Variables', 'Observaci', 'Tamaño', 'Memoria', and 'Ordenado'.

# Análisis factorial exploratorio

Estimamos con la cantidad de factores sugeridas por el Screeplot (3)

factor per1 per2 per3 per4 per5 per6 per7 per8 per9 per10 , factor (3)  
rotate, oblique promax

Rotated factor loadings (pattern matrix) and unique variances

Variable	Factor1	Factor2	Factor3	Uniqueness
per1	0.5204	-0.0888	0.0708	0.7153
per2	0.0453	0.0108	0.5062	0.7404
per3	0.5174	-0.0495	0.0917	0.7228
per4	0.0483	0.1923	0.4299	0.7404
per5	0.4631	-0.1201	0.1222	0.7589
per6	0.0841	0.5081	-0.0595	0.7503
per7	0.5684	0.1296	-0.0687	0.6668
per8	-0.1000	0.3850	0.2312	0.7414
per9	0.4222	0.1307	-0.1940	0.7812
per10	-0.0262	0.4652	0.1396	0.7319

# Análisis factorial exploratorio

Analizamos los ítems y procedemos con la creación de variables

Usted se considera una persona que es:	
PER1.	Sociable y activa ✓
PER2.	Una persona crítica y peleonera
PER3.	Una persona confiable y disciplinada
PER4.	Una persona ansiosa y fácil de molestarse
PER5.	Una persona abierta a nuevas experiencias e intelectual
PER6.	Una persona callada y tímida
PER7.	Una persona generosa y cariñosa
PER8.	Una persona desorganizada y descuidada
PER9.	Una persona calmada y emocionalmente estable
PER10.	Una persona poco creativa y con poca imaginación

# Análisis factorial exploratorio

Analizamos los ítems y procedemos con la creación de variables

Usted se considera una persona que es:	
PER1.	Sociable y activa
PER2.	Una persona crítica y peleonera
PER3.	Una persona confiable y disciplinada
PER4.	Una persona ansiosa y fácil de molestarse
PER5.	Una persona abierta a nuevas experiencias e intelectual
PER6.	Una persona callada y tímida ✓
PER7.	Una persona generosa y cariñosa
PER8.	Una persona desorganizada y descuidada ✓
PER9.	Una persona calmada y emocionalmente estable
PER10.	Una persona poco creativa y con poca imaginación ✓

# Análisis factorial exploratorio

Analizamos los ítems y procedemos con la creación de variables

Usted se considera una persona que es:	
PER1.	Sociable y activa
PER2.	Una persona crítica y peleonera ✓
PER3.	Una persona confiable y disciplinada
PER4.	Una persona ansiosa y fácil de molestar ✓
PER5.	Una persona abierta a nuevas experiencias e intelectual
PER6.	Una persona callada y tímida
PER7.	Una persona generosa y cariñosa
PER8.	Una persona desorganizada y descuidada
PER9.	Una persona calmada y emocionalmente estable
PER10.	Una persona poco creativa y con poca imaginación

Muchas gracias

