

# Ánalysis de las encuestas de hogares con R

Cápítulos: Procesamiento longitudinal de las encuestas rotativas

CEPAL - Unidad de Estadísticas Sociales

# Tabla de contenidos I

Procesamiento longitudinal de las encuestas rotativas

Análisis posibles con datos longitudinales

Procesamiento longitudinal de las encuestas rotativas

# Introducción

¿Por qué pensar en procesamiento longitudinal?

- ▶ Algunas oficinas requieren estadísticas de seguimiento continuo.
- ▶ Se aprovecha el esquema rotativo para generar información longitudinal.
- ▶ Requiere estructura de ponderación específica.

¿Qué es una encuesta longitudinal?

- ▶ Recolecta información sobre los mismos elementos en múltiples momentos.
- ▶ Contrasta con levantamientos transversales.
- ▶ Un ejemplo: esquema rotativo 4(1)0 permite seguimiento anual de 25%.

# Introducción

*Según Lynn (2009), una encuesta longitudinal observa los mismos elementos a lo largo del tiempo. Muchas encuestas rotativas, como las de empleo, se pueden convertir en longitudinales si se estructura adecuadamente el seguimiento.*

# Estimación del cambio y la varianza

- ▶ El foco está en estimar cambios entre periodos consecutivos.
- ▶ Es necesario calcular:
  - ▶ Varianza del periodo 1
  - ▶ Varianza del periodo 2
  - ▶ Correlación entre ambos
- ▶ Esos elementos afectan CV y tamaño de muestra.

*“Uno de los retos metodológicos clave es que las muestras no son independientes. Para estimar cambios correctamente, debemos considerar la varianza en cada ronda y la correlación entre observaciones repetidas.”*

Análisis posibles con datos longitudinales

# Caracterización de transiciones individuales

- ▶ Identificación de hogares/personas que cambian de estatus.
- ▶ Análisis de las características de quienes *entran o salen* de situaciones como la pobreza extrema, incluso sin cambios netos agregados.

## *Estabilidad e inestabilidad en el tiempo*

- ▶ Seguimiento prolongado permite detectar *trayectorias persistentes o fluctuaciones*.
- ▶ Comprensión más profunda de los factores que **explican la permanencia en condiciones como la pobreza extrema**.



# Duración, eventos e impactos

## Caracterización de eventos y duración

- ▶ Estudio de la *duración de estados*: cuánto tiempo se permanece desempleado, inactivo, fuera del sistema educativo, etc.
- ▶ Posibilidad de construir *indicadores de duración y persistencia*.

## Evaluación de impactos y relaciones causales

- ▶ Estimación del **efecto de intervenciones** o choques externos (ej. COVID-19) sobre fenómenos como la **desocupación**.

# Diseño de paneles rotativos en encuestas de hogares

- ▶ En América Latina, varias encuestas de hogares incorporan esquemas de panel rotativo que permiten la observación repetida de una misma unidad de análisis.
- ▶ Este diseño busca:
  - ▶ Capturar dinámicas intra-hogar e interpersonales a lo largo del tiempo.
  - ▶ Generar estimaciones robustas sobre cambios de estado (e.g., ocupación inactividad).
- ▶ La dimensión longitudinal se configura sobre los hogares que **respondieron efectivamente en más de un periodo**, permitiendo análisis más allá de los cortes transversales.

## Traslapes muestrales en un esquema 4(0)1

- ▶ Una encuesta con diseño **4(0)1** realiza cuatro observaciones trimestrales consecutivas por vivienda antes de su rotación definitiva.
- ▶ La rotación de paneles genera **traslapes sistemáticos** entre periodos consecutivos:
  - ▶ **T1 vs T2** → **75%** de hogares compartidos.
  - ▶ **T1 vs T3** → **50%**
  - ▶ **T1 vs T4** → **25%**
  - ▶ **T1 vs T5** → **0%** (panel completamente renovado).
- ▶ Esta propiedad escalonada permite construir **bases longitudinales de corto, mediano y largo plazo**, dependiendo del objetivo analítico.

# Traslapes muestrales en un esquema 4(0)1

Tabla 1: *Rotación de páneles para un diseño 4(0)1.*

Trimestre	Panel 1	Panel 2	Panel 3	Panel 4
T1	$a_1$	$b_1$	$c_1$	$d_1$
T2	$b_1$	$c_1$	$d_1$	$a_2$
T3	$c_1$	$d_1$	$a_2$	$b_2$
T4	$d_1$	$a_2$	$b_2$	$c_2$
T5	$a_2$	$b_2$	$c_2$	$d_2$
T6	$b_2$	$c_2$	$d_2$	$a_3$

# Construcción de esquemas longitudinales en paneles rotativos

La figura ilustra tres esquemas longitudinales que pueden derivarse del diseño rotativo:

1. **Bimestral (T1–T2):**
2. **Trimestral extendido (T1–T3):**
3. **Anual (T1–T4):**

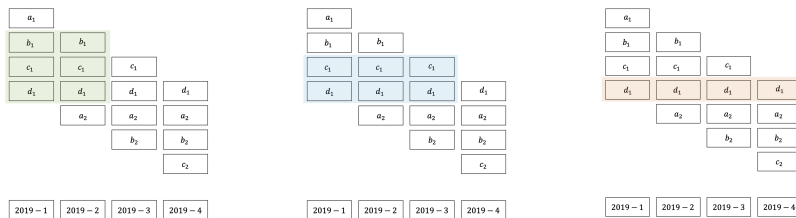


Figura 1: Tres escenarios longitudinales en un esquema rotativo 4(0)1.

## Efectos del COVID-19 en el esquema rotativo – Año 2020

- ▶ El año 2020 representó un quiebre en la operatividad normal de los levantamientos debido a la emergencia sanitaria provocada por el COVID-19.
- ▶ Las *restricciones de movilidad* obligaron a:
  - ▶ Cambiar el modo de recolección a entrevistas **telefónicas**, reduciendo la cobertura y la tasa de respuesta efectiva.
  - ▶ *Repetir el diseño muestral* del primer trimestre en el tercer trimestre, lo que alteró la lógica original de rotación 4(0)1.

## Esquema observado:

Año	Trimestre	Panel 1	Panel 2	Panel 3	Panel 4
2020	T1	$a_1$	$b_1$	$c_1$	$d_1$
	T2	$b_1$	$c_1$	$d_1$	$a_2$
	T3	$b_1$	$c_1$	$d_1$	$a_2$
	T4	$c_1$	$d_1$	$a_2$	$b_2$

### Implicaciones en el traslape muestral:

- ▶ **T2 vs T3:** 100% de traslape.
- ▶ **T1 vs T3:** 75% de traslape.
- ▶ **T1 vs T4:** 50% de traslape.

# Cargue de base de datos y librerías

```
library(printr)      # Mejora la presentación de tablas en documentos RMarkdown
library(tidyverse)   # Conjunto de paquetes para manipulación y
                     # visualización de datos
library(tidyr)       # Manejo de estructuras anchas/largas
library(pROC)        # Curvas ROC
library(survey)      # Análisis de encuestas con diseño muestral complejo

# Carga de la base de datos a nivel de personas
base_personas <- readRDS(file.path(input, "base_personas.rds")) %>%
  ungroup() # Elimina cualquier agrupamiento previo

# Creación de la base a nivel de hogares, extrayendo variables
# clave sin duplicados

base_hogares <- base_personas %>%
  distinct(upm, trimestre, id_hogar, fep)

# Visualización preliminar de la base de hogares
head(base_hogares, 10)
```



## Cargue de base de datos y librerías

upm	trimestre	id_hogar	fep
1100100006	t1	262	19
1100100006	t2	262	19
1100100006	t1	265	16
1100100006	t2	265	16
1100100006	t1	277	16
1100100006	t2	277	16
1100100006	t1	288	19
1100100006	t2	288	19
1100100006	t1	289	30
1100100006	t2	289	30

## Muestra de UPMs por trimestre

```
# Número de hogares por trimestre
base_hogares %>%
  group_by(trimestre) %>%
  tally(name = "hogares")
```

trimestre	hogares
t1	5000
t2	4846

```
# Número de UPMs únicas por trimestre
base_hogares %>%
  distinct(trimestre, upm) %>%
  group_by(trimestre) %>%
  tally(name = "upm")
```

trimestre	upm
t1	500
t2	495

# Número de UPMs en el traslape

Paso 1: identificar hogares que aparecen en ambos trimestres

```
# Identifica hogares que aparecen exactamente en dos trimestres
hogares_ambos <- base_hogares %>%
  group_by(id_hogar) %>%
  count() %>% filter(n == 2) %>% # Aparecen en dos trimestres
  pull(id_hogar)

# Extrae los pesos del trimestre 1 para los hogares que serán comparados
base_t1 <- base_hogares %>%
  filter(trimestre == "t1") %>% select(id_hogar, fep_t1 = fep)

# Número total de hogares con traslape en dos trimestres
length(hogares_ambos)

[1] 3627
```

¡Gracias!

*Email:* [andres.gutierrez@cepal.org](mailto:andres.gutierrez@cepal.org)

# Referencias

Lynn, P. 2009. *Methodology of longitudinal surveys*. Wiley series en survey methodology. Wiley.