República del Ecuador

Pesos replicados en la ENEMDU

Instituto Nacional de Estadística y Censos (INEC) - Ecuador

Septiembre-2021

www.ecuadorencifras.gob.ec

INEC Buenas cifras, mejores vidas

Buenas cifras, mejores vidas

01.

ENEMDU

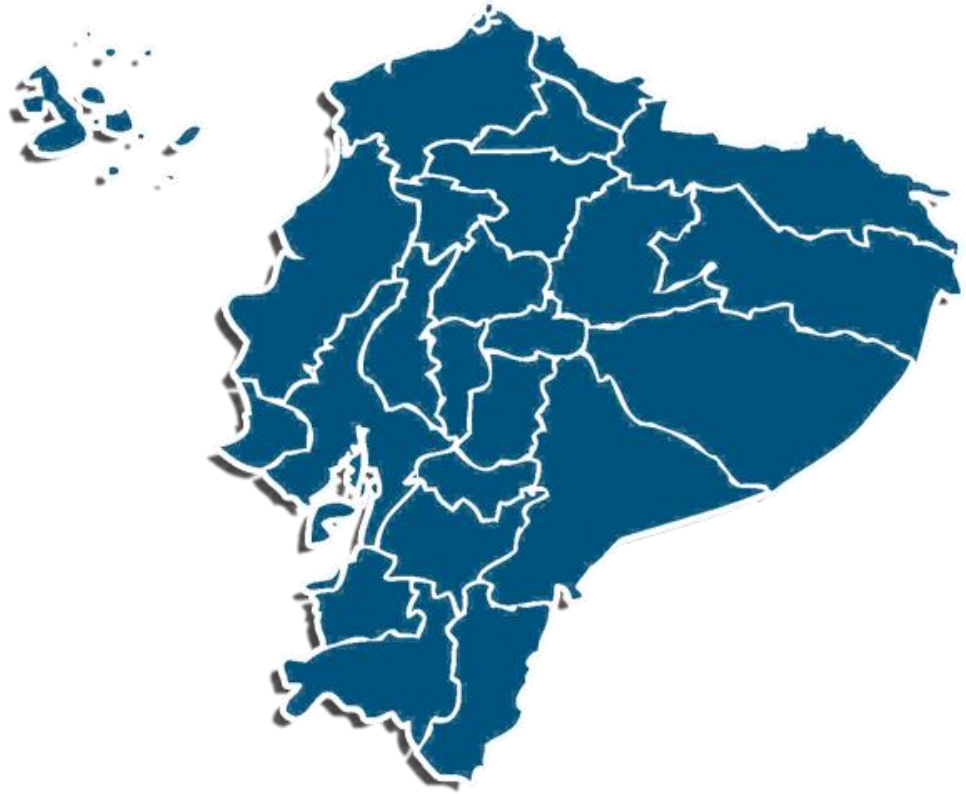
Encuesta Nacional de Empleo, Desempleo y Subempleo

Encuesta Nacional de Empleo

Ficha técnica

Ficha técnica de la encuesta

Tipo de encuesta	Longitudinal / panel de viviendas - Continua
Periodicidad	Mensual Trimestral Anual
Cobertura geográfica	Nacional Urb/Rur 5 ciudades principales Provincial
Población objetivo	Personas de 5 años y más (informante directo y calificado).

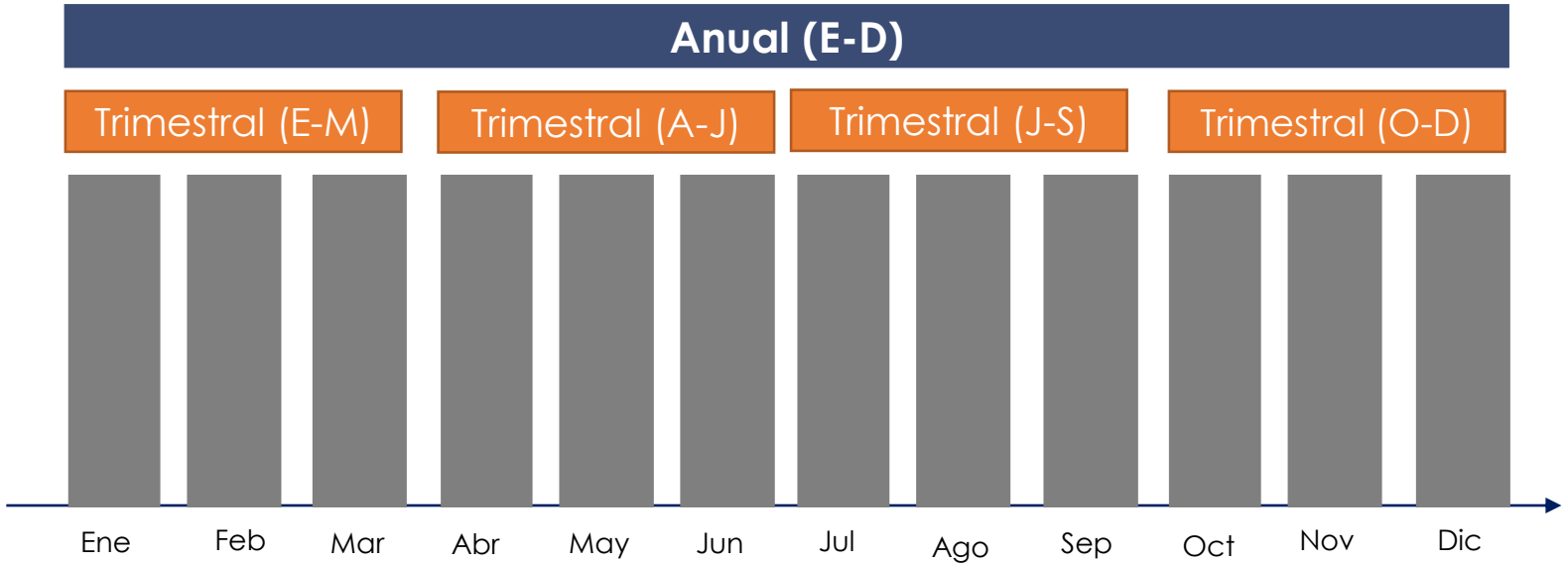


La ENEMDU es una encuesta por muestreo probabilístico, cuyo propósito principal es la medición y seguimiento del empleo, desempleo y la caracterización del mercado de trabajo, que permite conocer la actividad económica y las fuentes de ingresos de la población.

Encuesta Nacional de Empleo

Estimaciones

Una encuesta continua, con levantamientos mensuales ininterrumpidos en todo el año, permite brindar estimaciones **de carácter trimestral y anual.**



Encuesta Nacional de Empleo

Acumulación

Producción Estadística puntual		Acumulación de muestra
Estimadores puntuales		Estimadores acumulados
Situación **en un momento** del tiempo		Situación **durante** el tiempo
Ej: A dic-2020, la tasa de subempleo a nivel nacional fue de 16,4%.		**Ej: En 2020**, la tasa de subempleo a nivel nacional fue de 16,9%.

Encuesta Nacional de Empleo

Rotación de la muestra

	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
Panel	**A11**	**E11**	**I11**	**A11**	**E11**	**I11**	**A12**	**E12**	**I12**	**A12**	**E12**	**I12**
Viviendas	2.254	2.254	2.254	2.254	2.254	2.254	2.254	2.254	2.254	2.254	2.254	2.254
Panel	**B11**	**F11**	**J11**	**B12**	**F12**	**J12**	**B12**	**F12**	**J12**	**B11**	**F11**	**J11**
Viviendas	2.254	2.254	2.254	2.254	2.254	2.254	2.254	2.254	2.254	2.254	2.254	2.254
Panel	**C11**	**G11**	**K11**	**C11**	**G11**	**K11**	**C12**	**G12**	**K12**	**C12**	**G12**	**K12**
Viviendas	2.254	2.254	2.254	2.254	2.254	2.254	2.254	2.254	2.254	2.254	2.254	2.254
Panel	**D11**	**H11**	**L11**	**D12**	**H12**	**L12**	**D12**	**H12**	**L12**	**D21**	**H21**	**L21**
Viviendas	2.254	2.254	2.254	2.254	2.254	2.254	2.254	2.254	2.254	2.254	2.254	2.254
Total Viviendas por trimestre	**27.048**			**27.048**			**27.048**			**27.048**		

Características de la ENEMDU:

- Encuesta que incorpora **paneles de rotación** con duración de dos años.
- Seguimiento **2-2-2**, es decir la mitad de la muestra se repetirá cada trimestre
- Cada UPM pertenece a un panel, el cual está constituido por siete viviendas

Encuesta Nacional de Empleo

Buenas cifras, mejores vidas

Principales indicadores

Indicadores de mercado laboral

- Participación bruta
- Participación global
- Desempleo
- Empleo bruto
- Empleo global
- Empleo adecuado
- Subempleo
- Empleo no pleno
- Empleo no remunerado
- Empleo en el sector formal e informal

Indicadores de pobreza

- Pobreza por ingresos
- Extrema pobreza por ingresos
- Pobreza por Necesidades Básicas Insatisfechas (NBI)
- Pobreza Multidimensional

Indicadores de desigualdad

- Índice de Gini

ENEMDU

Estimación de la varianza

La estimación de la varianza que se realiza para los indicadores de mercado laboral, pobreza y desigualdad que se obtienen a partir de la ENEMDU, es calculada mediante:

Linealización → Técnica del último conglomerado

R (librería srvyr)

```
plan <- base %>%
  as_survey_design(ids = id_upm,
                   strata = estrato,
                   weights = fexp,
                   nest = T)
options(survey.lonely.psu="certainty")
```

STATA

```
svyset id_upm [iw=fexp], strata(estrato)
vce(linearized) singleunit(certainty)
```

Buenas cifras, mejores vidas

02.

Pesos replicados

Ejercicios y simulaciones realizadas

Réplicas

Definición y ventajas

Alternativas para estimar los errores estándar:

- Fórmulas exactas
- Linealización
- Réplicas

Réplicas:

- Las réplicas corresponden a una técnica que permite estimar los errores estándar de cualquier indicador (sin importar su complejidad).

- Los usuarios pueden estimar los errores de muestreo de forma automatizada, sin necesidad de aplicar complejas fórmulas matemáticas.

- No es indispensable conocer el diseño de muestreo que se utilizó en una encuesta.

- Protege la anonimidad de los respondientes.

Pesos replicados

Metodologías de cálculo

Jackknife

Técnica en la cual los ponderadores replicados se calculan retirando UPM en cada réplica. Se generan tantas réplicas como número de UPM existan.

- JK1
- JK2
- JKn
- Jackknife agrupado

Réplicas repetidas balanceadas (BRR)

Técnica que se estableció para diseños muestrales en donde se seleccionan dos UPM por estrato, produce estimaciones eficientes para estimaciones lineales y no lineales (incluidos los percentiles).

- BRR estándar
- BRR con ajuste de Fay

Bootstrap

Produce estimaciones eficientes del error estándar para indicadores lineales y no lineales. Es aplicable cuando se tiene dos o más UPM por estrato. En cada estrato se selecciona una submuestra de UPM con reemplazamiento.

- Bootstrap clásico
- Sub bootstrap
- Mrb bootstrap

Pesos replicados en la ENEMDU

Buenas cifras, mejores vidas

Ejercicios y simulaciones realizadas

1. Cálculo del número óptimo de réplicas

- Se calculó el número óptimo de réplicas, considerando la metodología del Bootstrap (bootstrap clásico, subbootstrap y mrbbootstrap).
- Se realizó este ejercicios para estimaciones de totales, tasas e índice de Gini.
- Se calcularon las estimaciones del error estándar (como medida absoluta) y del coeficiente de variación (como medida relativa).

2. Calibración de las réplicas

- Análisis de la calibración en las réplicas obtenidas a partir de la metodología Bootstrap (bootstrap clásico, subbootstrap y mrbbootstrap).
- Calibración de las réplicas con información auxiliar.

3. Comparación de los errores estándar

- Se comparó las estimaciones de los errores estándar y de los coeficientes de variación obtenidos a través de la linealización y de las diferentes metodologías de cálculo de pesos replicados.

Cálculo del número óptimo de réplicas

Buenas cifras, mejores vidas

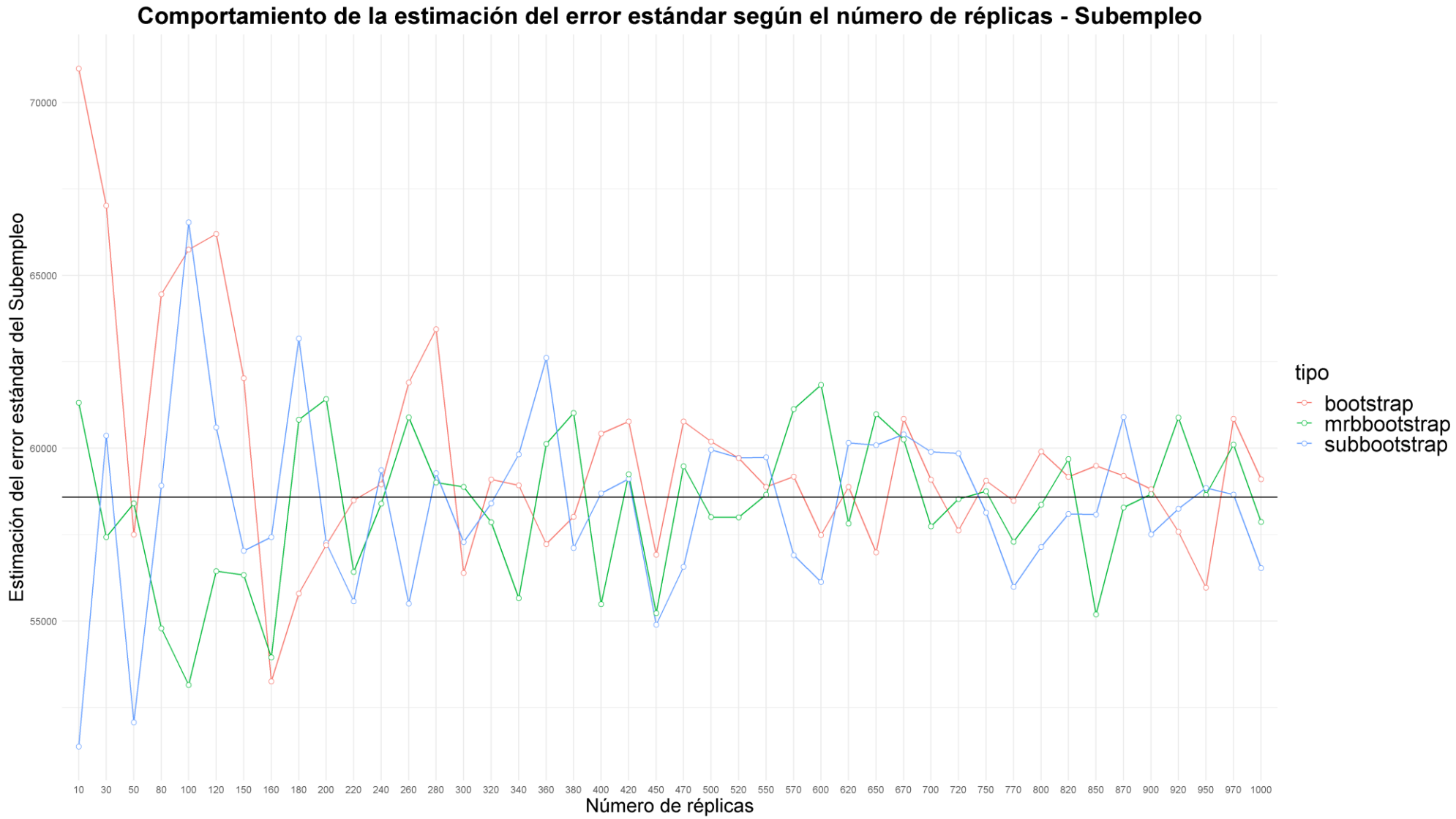
El valor del número óptimo de réplicas se obtuvo identificando la estimación del error estándar obtenida con las réplicas más cercana o más aproximada al valor del error estándar obtenido con la técnica de linealización.

Subbootstrap	**Totales**						
	Población	PEA	Adec	Desem	Sub	Pobreza	Ext. Pobreza
	850	650	620	240	970	100	900
	Tasas						
	Ing. Medio	Adec	Desem	Sub	Pobreza	Ext. Pobreza	
	820	970	670	120	870	900	
	Gini						
	770						
Mrbootstrap	**Totales**						
	Población	PEA	Adec	Desem	Sub	Pobreza	Ext. Pobreza
	570	850	1000	380	720	920	570
	Tasas						
	Ing. Medio	Adec	Desem	Sub	Pobreza	Ext. Pobreza	
	620	700	550	470	870	420	
	Gini						
	670						
Bootstrap	**Totales**						
	Población	PEA	Adec	Desem	Sub	Pobreza	Ext. Pobreza
	200	160	720	120	220	260	650
	Tasas						
	Ing. Medio	Adec	Desem	Sub	Pobreza	Ext. Pobreza	
	80	260	120	720	500	750	
	Gini						
	800						

Cálculo del número óptimo de réplicas

Buenas cifras, mejores vidas

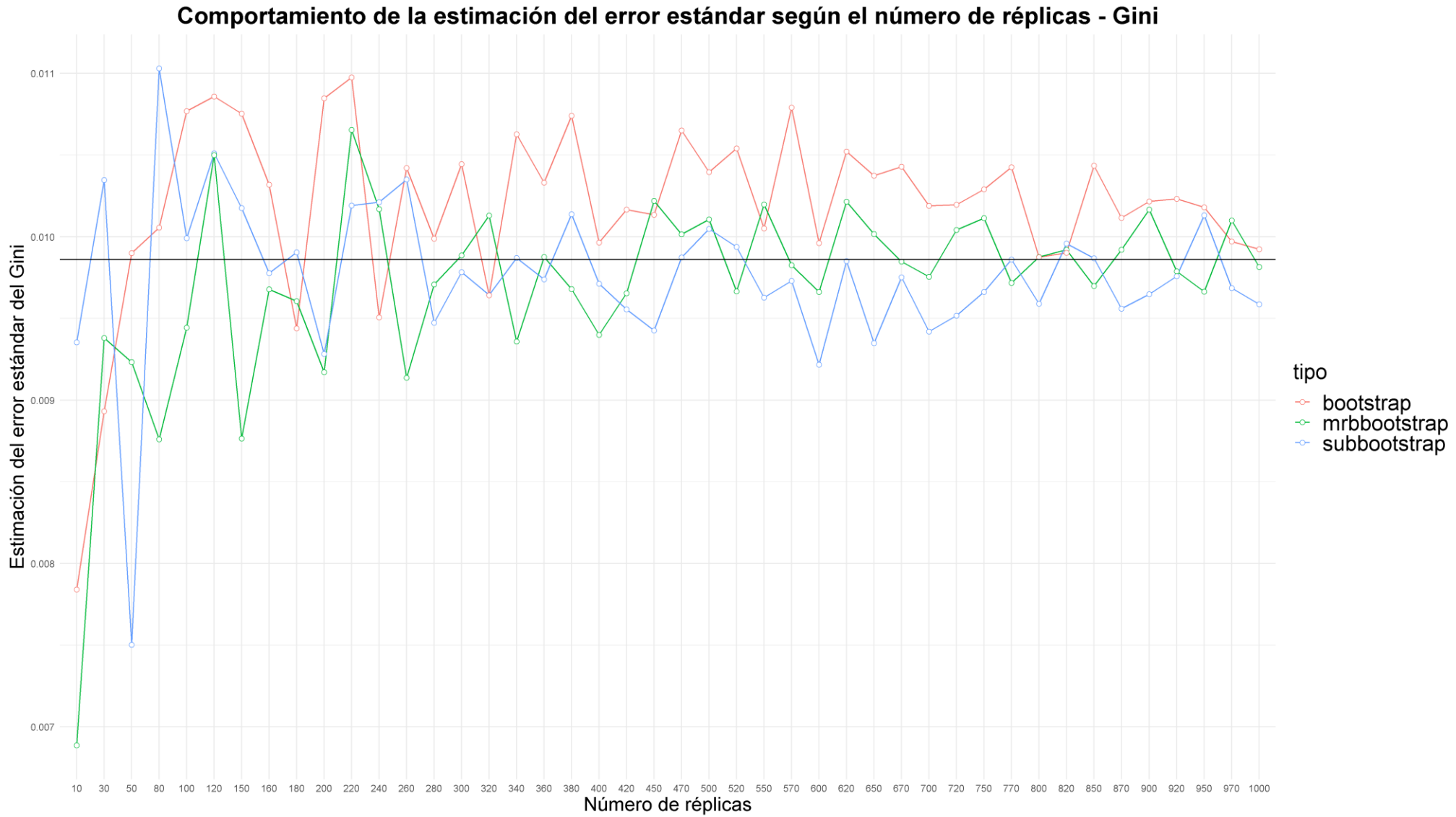
Total de subempleados



Cálculo del número óptimo de réplicas

Buenas cifras, mejores vidas

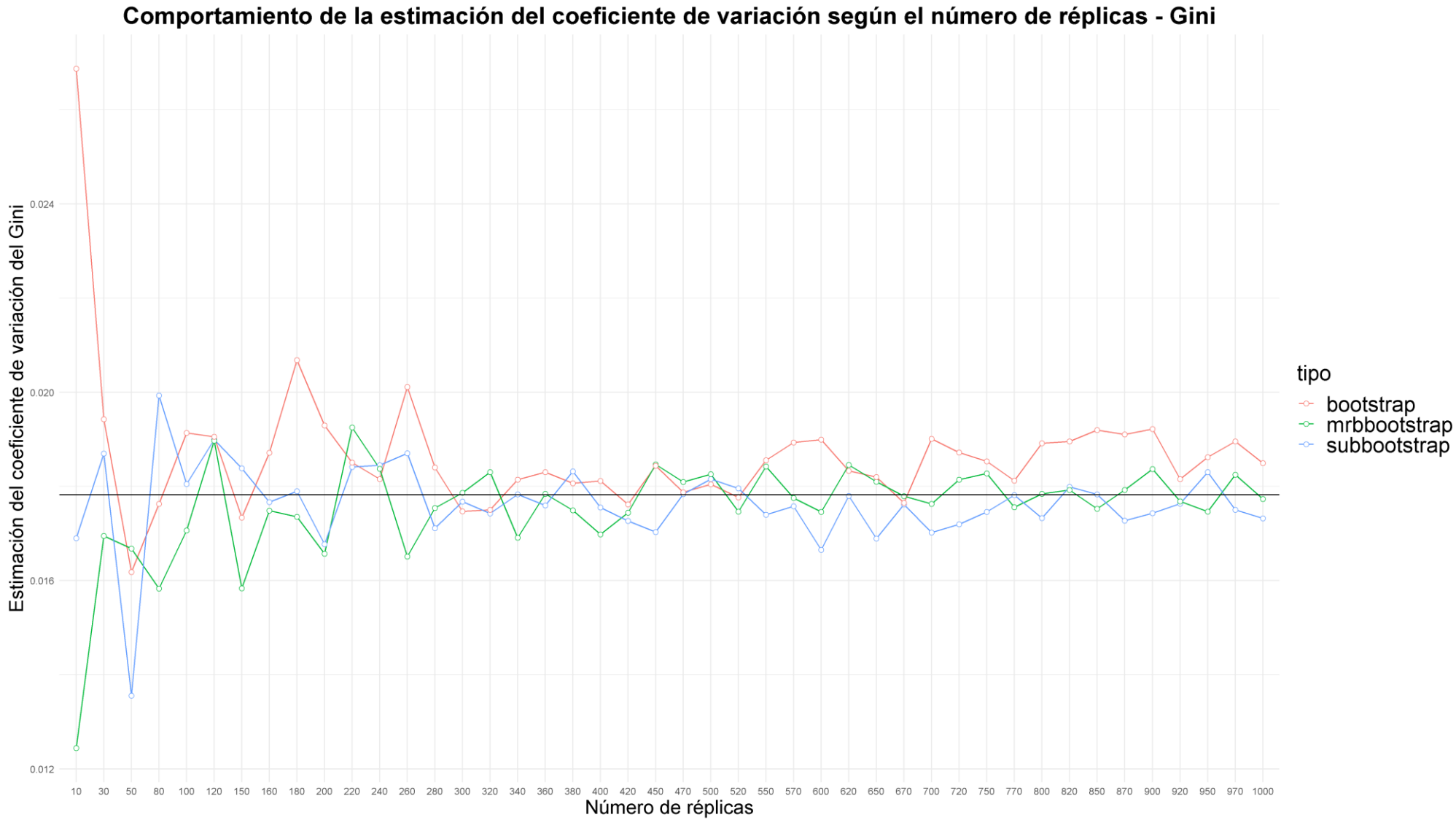
Índice de Gini – error estándar



Cálculo del número óptimo de réplicas

Buenas cifras, mejores vidas

Índice de Gini – coeficiente de variación



Cálculo del número óptimo de réplicas

Buenas cifras, mejores vidas

Resultados de los ejercicios y simulaciones realizadas

1. Cálculo del número óptimo de réplicas

- Para la mayoría de indicadores (totales, tasas y Gini), la metodología del bootstrap clásico producía estimaciones del error estándar más alejadas del valor estimado del error estándar obtenido con la técnica de la linealización.

- Las metodologías de construcción de pesos replicados (subbootstrap y mrbbootstrap) generaban estimaciones del error estándar más cercanos y aproximados al valor estimado del error estándar obtenido con la técnica de la linealización.

- Para la mayoría de indicadores, a partir de las 200 réplicas, se observaba un comportamiento homogéneo (con menor dispersión), del error estándar y del coeficiente de variación, por lo que se determinó que 200 era el valor del número óptimo de réplicas.

Calibración de las réplicas

Buenas cifras, mejores vidas

Resultados de los ejercicios y simulaciones realizadas

Se analizó las réplicas obtenidas a partir de la calibración de factores de expansión en base a dos librerías:

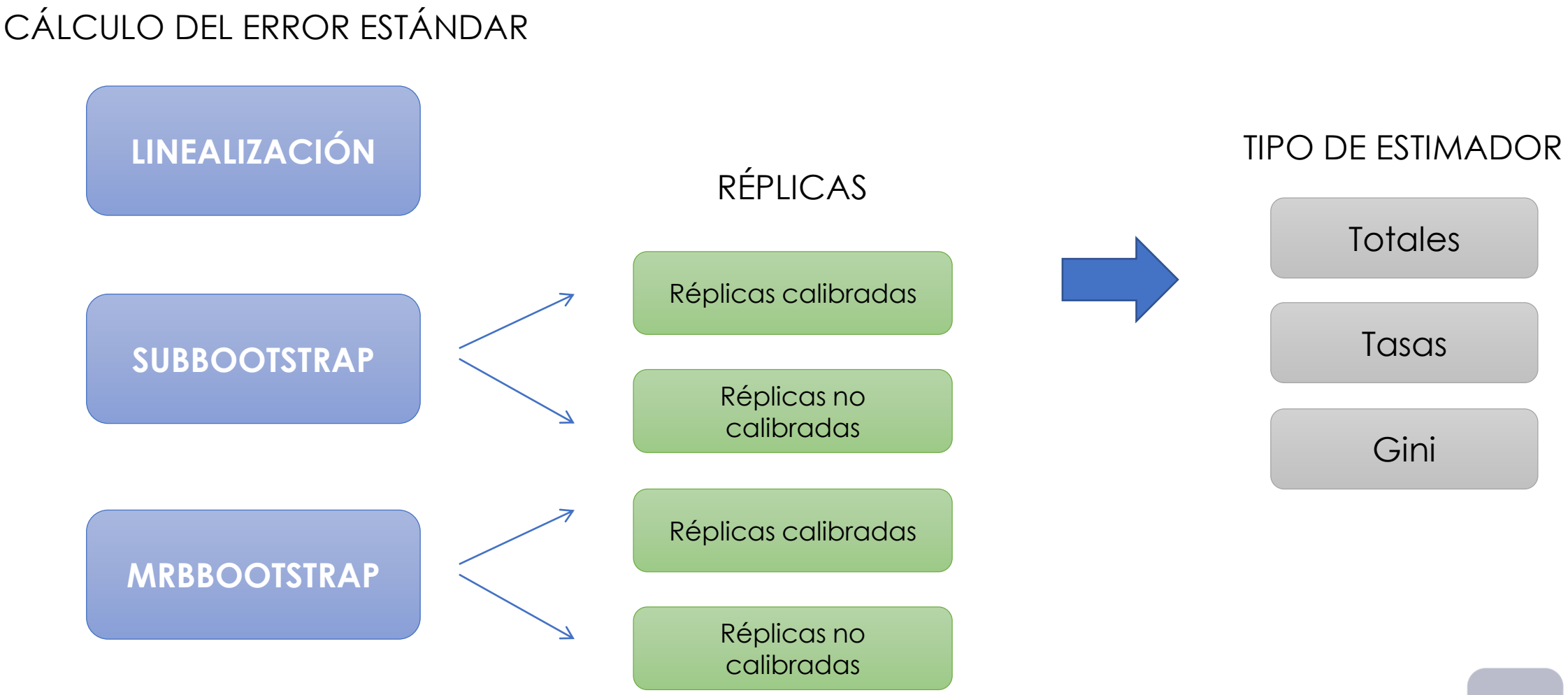
- Library(sampling) – comando (calib)
- Library(survey) -- comando (calibrate)

Resultado: el cálculo de pesos replicados, que parten de objetos calibrados a través de "sampling" o de "survey" producen réplicas que no están calibradas, es decir, no reproducen con exactitud la información auxiliar.

Ante este resultado obtenido, se procedió a calibrar las réplicas con información auxiliar (proyecciones de población), de la manera tradicional, procurando que las 200 réplicas reproduzcan con exactitud las proyecciones poblacionales de un mes de referencia (enero de 2021).

Pesos replicados en la ENEMDU

Comparación de los errores estándar y coeficientes de variación



Comparación – Error estándar

Población

Comparación – Error estándar

Tasa de desempleo

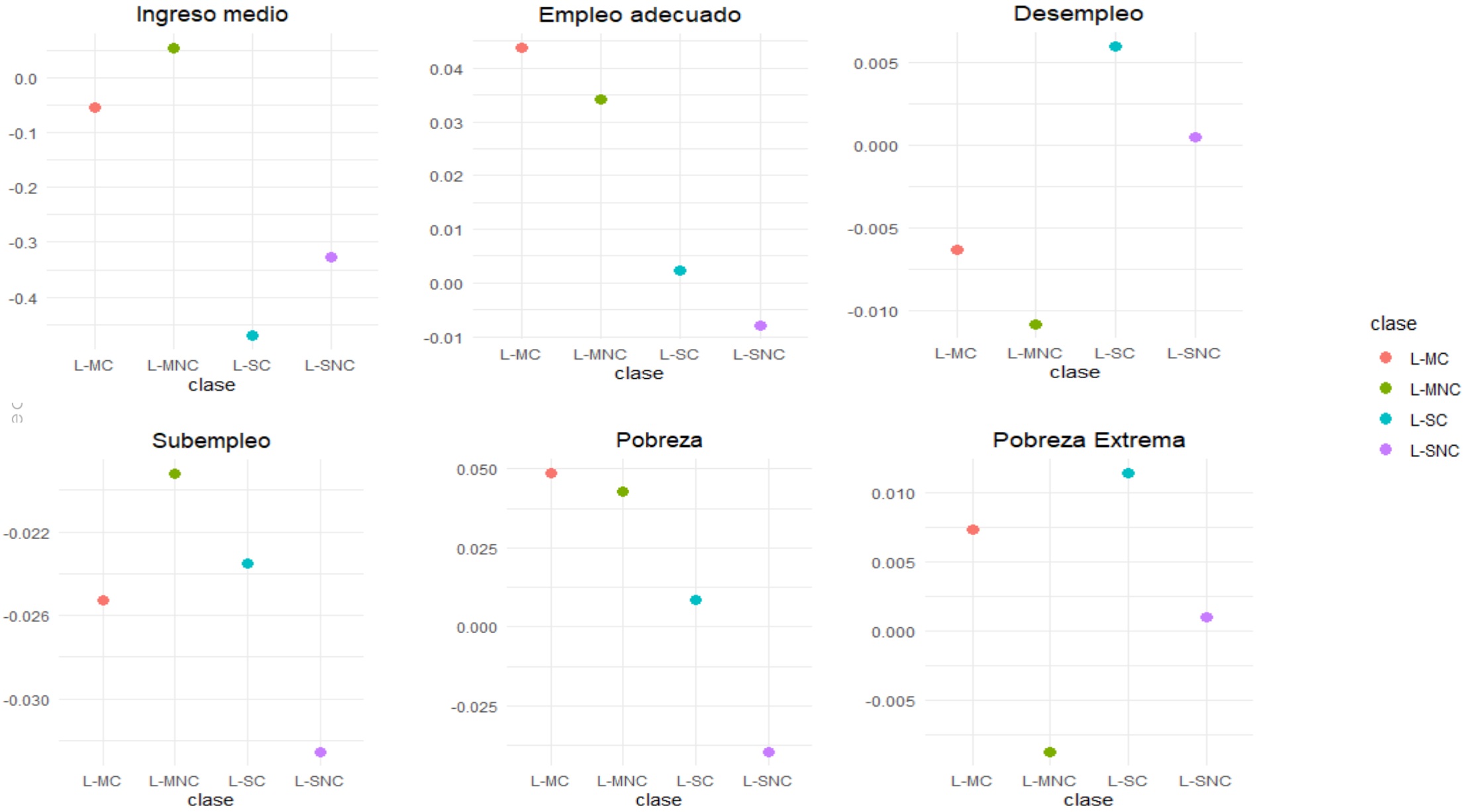
Comparación – Coeficiente de variación

Buenas cifras, mejores vidas

Tasa de pobreza por ingresos

Diferencias en estimadores de tasas

Buenas cifras, mejores vidas



Comparación de metodologías

Buenas cifras, mejores vidas

Resultados de los ejercicios y simulaciones realizadas

- Para la mayoría de indicadores, las metodologías subbootstrap y mrbbootstrap generaban estimaciones del error estándar cercanos o aproximados a los valores calculados a partir de la linealización.

- Las estimaciones del error estándar que partían de réplicas calibradas, eran más eficientes (error estándar más bajo) en comparación con las obtenidas en base a réplicas no calibradas.

- Por lo general, las estimaciones del error estándar que partían de réplicas calibradas, eran más eficientes con respecto a las obtenidas a partir de la técnica de linealización.

Pesos replicados en la ENEMDU

Aprendizaje y retos en el Ecuador

- Se utilizó la ENEMDU para la realización de estos ejercicios y simulaciones ya que es una encuesta multipropósito que reporta no sólo indicadores de mercado laboral, sino que reporta estimaciones de pobreza y desigualdad cuyo cálculo de la varianza es complejo.

- La metodología de pesos replicados es una forma de cálculo de factores de expansión que presenta algunas ventajas, por lo que se podría aplicar en otras operaciones estadísticas que se realizan en el país.

- Antes de la implementación, es necesario socializar más ampliamente estas metodologías tantos con los usuarios internos y externos de la información.

- La serie de ejercicios y simulaciones que se están llevando a cabo con acompañamiento técnico de la CEPAL todavía no está terminada, de tal manera que se podría seguir encontrando hallazgos y resultados interesantes referente a esta temática.

Pesos replicados en la ENEMDU

Correos institucionales y referencias

CORREOS INSTITUCIONALES

- Ing. Andrés Albán
DIRECTOR DE INFRAESTRUCTURA ESTADÍSTICA Y MUESTREO
Andres_Alban@inec.gob.ec

- Ing. Francisco Céspedes
JEFE DE LA GESTIÓN DE DISEÑO MUESTRAL
Francisco_Cespedes@inec.gob.ec

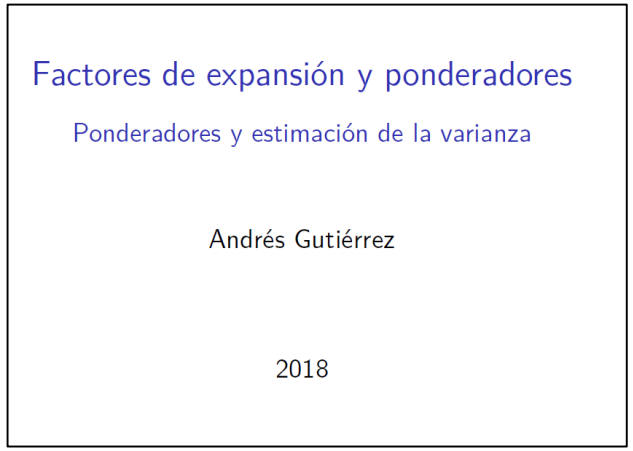
- Mat. Javier Núñez
MIEMBRO DE EQUIPO
Javier_Nunez@inec.gob.ec

- Ing. William Constante
MIEMBRO DE EQUIPO
William_Constante@inec.gob.ec

- Ing. Diana Simbaña
MIEMBRO DE EQUIPO
Diana_Simbana@inec.gob.ec

REFERENCIAS

Definiciones y teoría tomada del archivo .pdf proporcionado por CEPAL relacionado con las diferentes técnicas utilizadas para calcular los errores de muestreo.


Factores de expansión y ponderadores

Ponderadores y estimación de la varianza

Andrés Gutiérrez

2018

Buenas cifras,
mejores vidas

@ecuadorencifras @ecuadorencifras @InecEcuador t.me/ecuadorencifras INEC/Ecuador INECEcuador INEC Ecuador