

Taller

Uso de Encuestas de Hogares para la medición de los ODS sobre pobreza y desigualdad

**Aspectos específicos de la medición:
no respuesta en las variables del
ingreso**

CEPAL, División Estadística



UNITED NATIONS



ERRORES DE MEDICION

- Se clasifican en dos grandes grupos:
 - Errores asociados al muestreo
 - Errores ajenos al muestreo
- **Errores asociados al muestreo**
 - Se deben a que los parametros de la distribucion de una variable en la poblacion se estiman a partir de una muestra de esa poblacion
 - Los parametros quedan afectados por la variabilidad de la muestra
 - El error asociado al muestreo es una medida de esa variabilidad
- Reportar los errores asociados al muestreo junto con las estimaciones puntuales permite:
 - Mejorar la comprension de los fenomenos asociados a las estimaciones a partir de encuestas
 - Relativizando o afirmando la magnitud de las variaciones registradas
 - Estableciendo limites precisos a la confiabilidad que asignamos a una estimacion (en particular cuando se le solicita a la muestra niveles de desagregacion que no puede soportar)



UNITED NATIONS

ECLAC

ERRORES DE MEDICION

- Errores ajenos al muestreo
 - Son "democraticos". Afectan por igual tanto a las estimaciones con base en encuestas como con base a censos de la poblacion.
- Comprenden:
 - Defectos en la codificacion de respuestas
 - Defectos en el procesamiento de los datos
 - Defectos en la construccion de los instrumentos de recoleccion de la informacion
 - Defectos en la construccion de las bases de datos (limites en el largo con que se almacenan las variables en el archivo de datos)
 - Defectos en los marcos de muestreo (desactualizacion, viviendas mal indicadas, etc)
 - No contacto con los habitantes de la vivienda seleccionada
 - Respuestas incorrectas a las preguntas realizadas
 - Falta de respuesta a algunos items por desconocimiento o negativa
 - Rechazo total a realizar la entrevista o completar una parte
- **LOS ERRORES AJENOS AL MUESTREO PUEDEN SESGAR LAS ESTIMACIONES Y AFECTAN SU CONFIABILIDAD**



FORMAS DE CONTROLAR LOS ERRORES AJENOS AL MUESTREO

- Se basan en acciones de chequeo y control en todas las etapas del proceso estadístico. Comprenden, pero no se agotan en:
 - Manutención de un plantel de encuestadores permanentes
 - Diseño de los materiales y cursos de entrenamiento para el personal de campo y oficina
 - Diseño cuidadoso de los instrumentos de recolección de información
 - Diseño y prueba de las rutinas de trabajo en campo y oficina
 - Diseño y prueba con el tiempo suficiente de los programas de control de calidad, la conformación de las bases de datos y los algoritmos de cálculo
 - Etc. Etc. Etc.
- **TODO TIPO DE CONTROL ES POCO PARA ASEGURAR LA CALIDAD DEL PROCESO ESTADÍSTICO**



PROCEDIMIENTOS PARA LIDIAR CON LA NO RESPUESTA

- Se pueden agrupar en dos tipos de acciones: en campo para levantar el rechazo; y en oficina, para hacerle frente una vez que se produjo. En campo:
 - Rutinas de control del seguimiento dado a la no respuesta, como parte del proceso de levantamiento de los datos (seguimiento "epidemiologico")
 - Reentrevista por parte del supervisor y/o el jefe de campo
 - Contacto con el informante desde la oficina regional o central
 - Citacion a brindar la informacion, si el marco legal permite ejercer coercion para obtener la respuesta
- Una vez confirmada la no respuesta se puede:
 - No hacer nada
 - Eliminar las unidades que no responden
 - Imputar en forma completa la unidad (ej. Encuestas economicas, censos)
 - Imputar parcialmente la respuesta mediante modelos estadisticos
 - Reponderar los factores de expansion aplicando tecnicas de calibracion (mas o menos sofisticadas)



UNITED NATIONS

ECLAC

Algunas buenas prácticas para el tratamiento de la no respuesta

- Identificar la no respuesta por sus causas, lo que puede hacerse utilizando distintos códigos para los missing values (existen paquetes informáticos que permiten hacerlo sin utilizar valores numéricos)
- Distinguir la no respuesta de la respuesta “Cero” (por ejemplo en variables de ingreso)
- No imputar valores en la variable original sino en una derivada, que contenga o bien los valores imputados, o bien los imputados y los no imputados. Ambas variables, originales e imputadas, debieran figurar en la base de datos que se pone a disposición de los usuarios.
- Publicar junto con la metodología los detalles técnicos de la imputación (alcance, criterios utilizados, metodología de cálculo, resultados)



UNITED NATIONS

ECLAC

EJEMPLO DE APLICACIÓN DE ALGUNOS METODOS DE IMPUTACION DE NO RESPUESTA EN INGRESOS



UNITED NATIONS

ECLAC

Situación inicial: ytoh con aprox. 10% de valores missing

RECODE of ytoh (Ingreso total-Hogar)	Freq.	Percent	Cum.
Ingreso mayor que Cero	16,518	89.91	89.91
Ingreso Cero	1,854	10.09	100.00
Total	18,372	100.00	

. tab ingresohecero [fw = peso]

RECODE of ytoh (Ingreso total-Hogar)	Freq.	Percent	Cum.
Ingreso mayor que Cero	24,496,925	89.59	89.59
Ingreso Cero	2,847,905	10.41	100.00
Total	27,344,830	100.00	



UNITED NATIONS

ECLAC

Caso base: se calcula ingreso per cápita y pobreza utilizando solo las observaciones con respuesta

```
. summ ingcorte0 [fw = peso]
```

Variable	Obs	Mean	Std. Dev.	Min	Max
ingcorte0	24496925	5708.365	5975.931	35	215000

```
. tab pobreza0 [fw=peso]
```

RECODE of ingcorte0 (Ingreso pc del hogar sólo respondentes)	Freq.	Percent	Cum.
Pobre extremo	3,248,555	13.26	13.26
Pobre no extremo	5,725,836	23.37	36.63
Fuera de pobreza	15,522,534	63.37	100.00
Total	24,496,925	100.00	

```
. summ pobreza0 [fw = peso]
```

Variable	Obs	Mean	Std. Dev.	Min	Max
pobreza0	24496925	.3663477	.481806	0	1



UNITED NATIONS

ECLAC

Primer caso, se calcula ingreso per cápita y pobreza suponiendo que $y_{toh} = 0$ (se supone un valor)

```
. summ ingcorte1 [fw = peso]
```

Variable	Obs	Mean	Std. Dev.	Min	Max
ingcorte1	27344830	5113.851	5918.843	0	215000

```
. tab pobreza1 [fw=peso]
```

RECODE of ingcorte1 (Ingreso pc del hogar (missing = 0))	Freq.	Percent	Cum.
Pobre extremo	6,096,460	22.29	22.29
Pobre no extremo	5,725,836	20.94	43.23
Fuera de pobreza	15,522,534	56.77	100.00
Total	27,344,830	100.00	

```
. summ pobreza11 [fw = peso]
```

Variable	Obs	Mean	Std. Dev.	Min	Max
pobreza11	27344830	.4323412	.4954011	0	1



UNITED NATIONS

ECLAC

Segundo caso, se calcula ingreso per cápita y pobreza suponiendo que $y_{toh} = \text{mean}(y_{toh})$ (se toma el ingreso medio entre quienes responden)

```
. summ ingcorte2 [fw = peso] if ingcorte2 > 0
```

Variable	Obs	Mean	Std. Dev.	Min	Max
ingcorte2	27344830	5857.129	5846.762	35	215000

```
. tab pobreza2 [fw=peso]
```

RECODE of ingcorte2 (Ingreso del hogar (missing = media))	Freq.	Percent	Cum.
Pobre extremo	3,248,555	11.88	11.88
Pobre no extremo	5,812,629	21.26	33.14
Fuera de pobreza	18,283,646	66.86	100.00
Total	27,344,830	100.00	

```
. summ pobreza22 [fw = peso]
```

Variable	Obs	Mean	Std. Dev.	Min	Max
pobreza22	27344830	.3313674	.4707048	0	1



UNITED NATIONS

ECLAC

Tercer caso, se calcula ingreso per cápita y pobreza imputando ytoh por medias condicionales a partir de un conjunto de variables del hogar y el jefe

Casos a imputar = 1854

Variable	Obs	Mean	Std. Dev.	Min	Max
-----+----- imputacion					
ytoh	16518	18007.72	14774.9	70	460000

ytoh	16518	18007.72	14774.9	70	460000
------	-------	----------	---------	----	--------

Casos sin imputar = 193

Casos sin donar = 0

Casos imputados = 1661

. summ ytoh ytoh_MC [fw=peso]

Variable	Obs	Mean	Std. Dev.	Min	Max
-----+-----					
ytoh	27344830	17011.86	16260.94	0	460000
ytoh_MC	27344830	18840.8	15818.24	0	460000

. summ ingcorte3 [fw = peso]

Variable	Obs	Mean	Std. Dev.	Min	Max
-----+-----					
ingcorte3	27344830	5767.046	5993.357	0	215000

. tab pobreza3 [fw=peso]

RECODE of ingcorte3 (Ingreso del hogar (missing = medias condicionadas))	Freq.	Percent	Cum.
Pobre extremo	3,772,892	13.80	13.80
Pobre no extremo	6,025,128	22.03	35.83
Fuera de pobreza	17,546,810	64.17	100.00
-----+-----			
Total	27,344,830	100.00	

. summ pobreza33 [fw = peso]

Variable	Obs	Mean	Std. Dev.	Min	Max
-----+-----					
pobreza33	27344830	.3583134	.4795049	0	1



UNITED NATIONS

ECLAC

Cuarto caso, se calcula ingreso per cápita y pobreza imputando ytoh por hotdeck a partir de un conjunto de variables del hogar y el jefe

```
Casos totales =      18372
Casos a imputar =     1854
Numero de grupos =   3862
Casos sin imputar =    193
Casos imputados =    1661
```

```
.      summ ytoh ytoh_HD [fw=peso]
```

Variable	Obs	Mean	Std. Dev.	Min	Max
ytoh	27344830	17011.86	16260.94	0	460000
ytoh_HD	27344830	18835.4	16152.46	0	460000

```
. summ ingcorte4 [fw = peso]
```

Variable	Obs	Mean	Std. Dev.	Min	Max
ingcorte4	27344830	5762.167	6140.274	0	215000

```
. tab pobreza4 [fw=peso]
```

RECODE of ingcorte4 (Ingreso del hogar (missing = hotdeck))	Freq.	Percent	Cum.
Pobre extremo	3,927,938	14.36	14.36
Pobre no extremo	6,167,983	22.56	36.92
Fuera de pobreza	17,248,909	63.08	100.00
Total	27,344,830	100.00	

```
. summ pobreza44 [fw = peso]
```

Variable	Obs	Mean	Std. Dev.	Min	Max
pobreza44	27344830	.3692077	.4825903	0	1



Resumen de resultados

Variable	Obs	Mean	Std. Dev.	Min	Max
ingcorte0	24496925	5708.365	5975.931	35	215000
ingcorte1	27344830	5113.851	5918.843	0	215000
ingcorte2	27344830	5857.129	5846.762	35	215000
ingcorte3	27344830	5767.046	5993.357	0	215000
ingcorte4	27344830	5762.167	6140.274	0	215000

. summ pobreza00 pobreza11 pobreza22 pobreza33 pobreza44 [fw = peso]

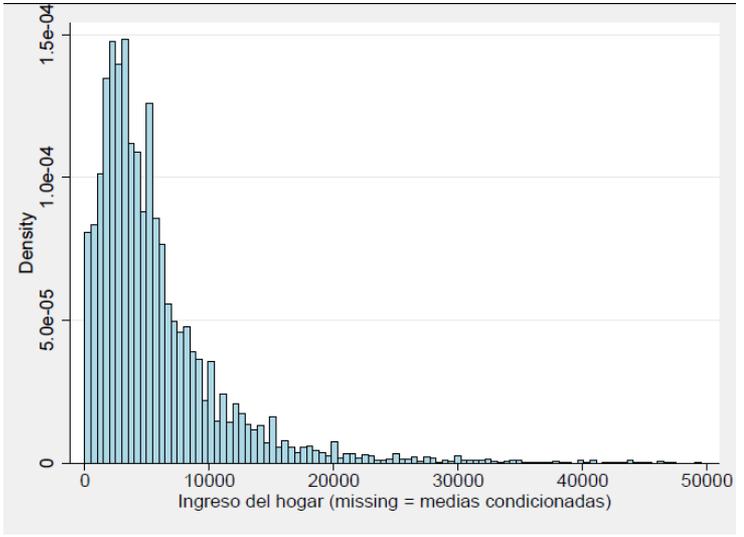
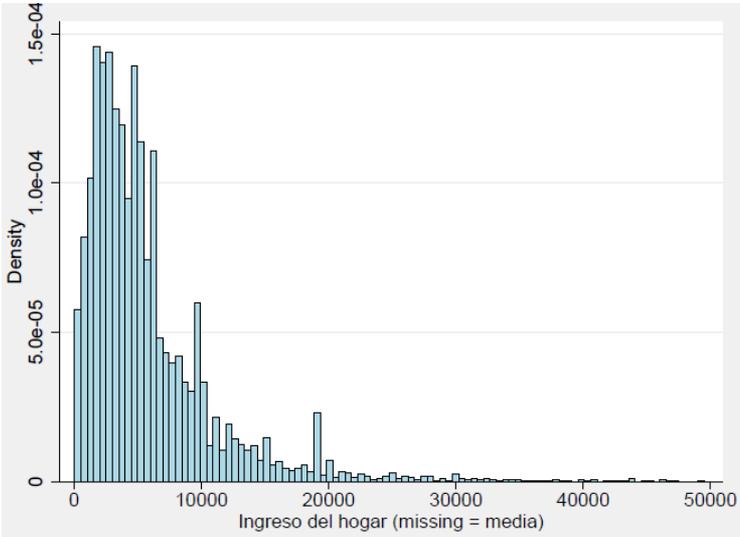
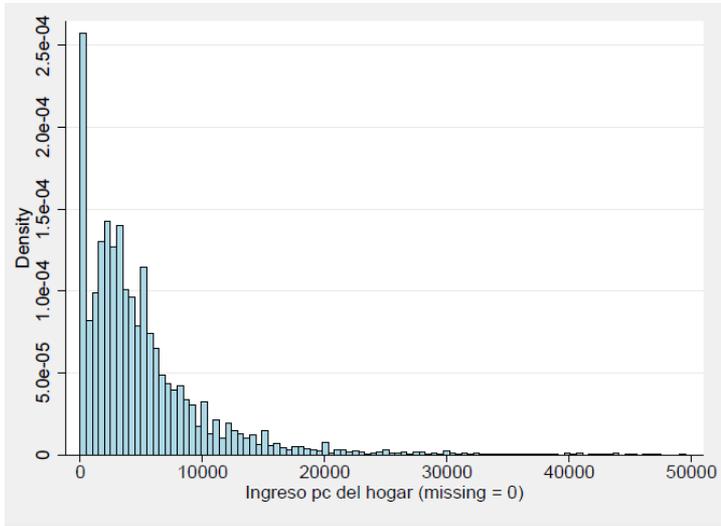
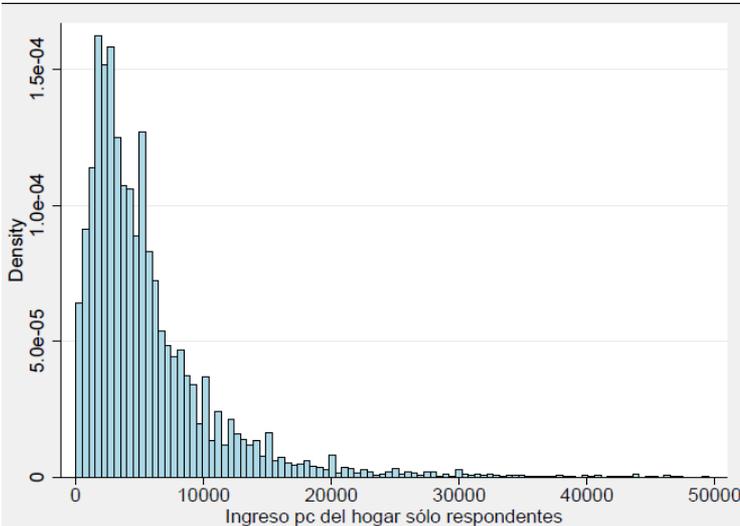
Variable	Obs	Mean	Std. Dev.	Min	Max
pobreza00	24496925	.3663477	.481806	0	1
pobreza11	27344830	.4323412	.4954011	0	1
pobreza22	27344830	.3313674	.4707048	0	1
pobreza33	27344830	.3583134	.4795049	0	1
pobreza44	27344830	.3692077	.4825903	0	1



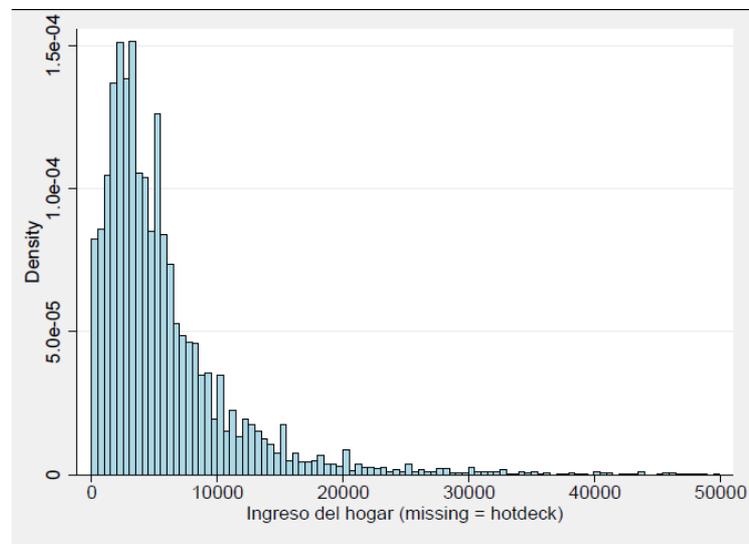
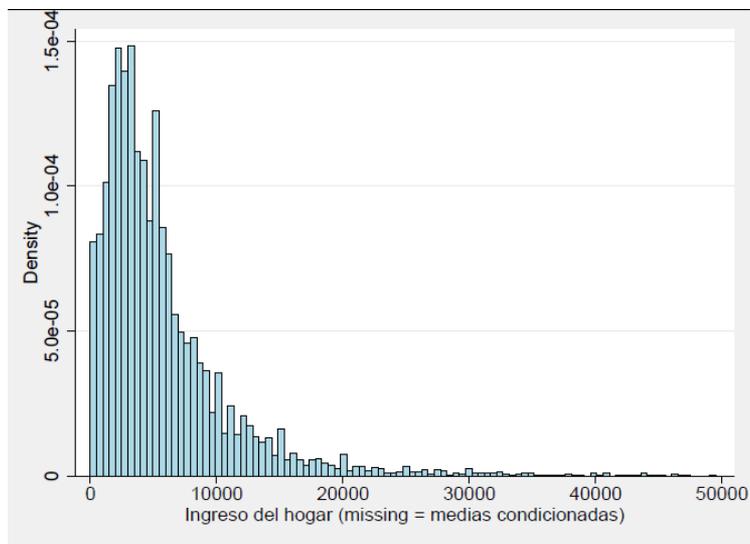
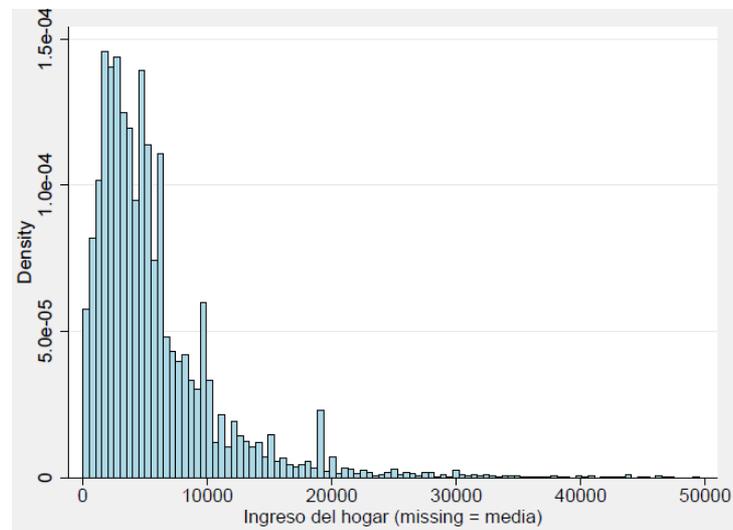
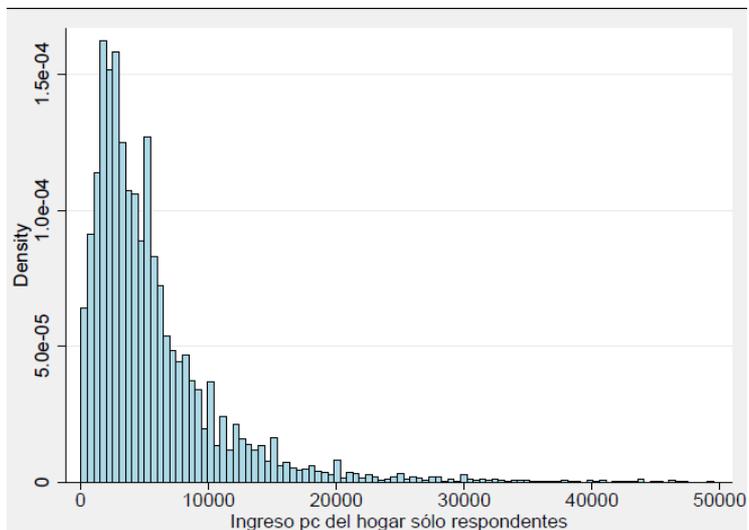
UNITED NATIONS

ECLAC

Análisis gráfico



Análisis gráfico



UNITED NATIONS

ECLAC

**Muchas gracias por
vuestra atención**



UNITED NATIONS

ECLAC

www.eclac.org