

Marcos muestrales imperfectos: herramientas para su tratamiento

Juan José Goyeneche

Instituto de Estadística
Facultad de Ciencias Económicas y de Administración
Universidad de la República
Uruguay

Modelos

Conceptos básicos

No existen marcos

Sí existen marcos

Conclusiones

not necessarily detract from its usefulness because all models
Essentially, all models are wrong, but some are useful.
ximate nature of the model must always be borne in mind

George Box en “Box, G., Draper, N.: (2007) Response Surfaces, Mixtures, and Ridge Analyses, 2nd ed. Wiley”, página 414.

Modelos perfectos

Marcos
imperfectos

Juan José
Goyeneche

Modelos

Conceptos básicos

No existen marcos

Sí existen marcos

Conclusiones



- ▶ $P(\text{sacar un } 3 \text{ en un dado}) = \dots$
- ▶ Si $P(3) = 0.17 \quad \rightsquigarrow \quad n = 48.020$

- Población, universo, población objetivo:

$$U = \{u_1, u_2, \dots, u_k, \dots, u_N\}$$

- ... que la representaremos:

$$U = \{1, 2, \dots, k, \dots, N\}$$

- ▶ Marco, marco muestral:
 - ▶ listado de elementos de U ,
 - ▶ listado de subconjuntos de elementos de U
- ▶ Muestra representativa

Modelos

Conceptos básicos

No existen marcos

Sí existen marcos

Conclusiones

REPUBLICA DE CHILE SERVICIO ELECTORAL		PADRON ELECTORAL DEFINITIVO - ELECCIONES PRESIDENCIAL, PARLAMENTARIAS Y DE CONSEJEROS REGIONALES 2013					PÁGINA	
REGION : DE COQUIMBO		COMUNA : LA SERENA					1 de 2	
PROVINCIA : ELQUI								
NOMBRE	C.IDENTIDAD	SEX	DOMICILIO ELECTORAL	CIRCUNSCRIPCION	MESA			
AARSAND DAN VIGGO	12.094.555-6	VAR	ANIBAL PINTO GARMENDIA 3870 DEPTO 23 ED. LA SILLA	LA SERENA	8			
ABALAN MARGALITA LIF ANTONIO	16.085.092-3	VAR	JOSÉ LUIS DABRE 1200 LA SERENA	LA SERENA	31			
ABACA LIZANA CRISTIAN MAURICIO	14.205.711-5	VAR	LOS AROMOS 3143 COMPANIA ALTA	LAS COMPANIAS	31			
ABAI AYARAY MARIO ADRIASO	14.116.533-K	VAR	CIA RIALA MONITAS 2525	LA SERENA	89			
ABALLAY NAIHOM ESTHER	17.339.625-3	MUJ	BEAUVASTA INTERIOR RING PARC. 7 S-15	LA SERENA	127			
ABAI AY CARVALJA AMALIA DEL ROSARIO	10.363.314-1	MUJ	P.I.11 FIOHUE 3480 VIAL A.F. ORIDA	LA SERENA	152			
ABALLAY CORTEZ MARGARET CAROLINE	9.802.535-9	MUJ	PARCELA 41 SITIO 15 BELLA VISTA	LA SERENA	142			
ABAIAY CORTEZ SHIRLEY MONICA	11.187.357-0	MUJ	PARCELA 6 SITIO 15 BELLA VISTA	LA SERENA	142			
ABALLAY FRUTIS MARCO PASQUAL	7.374.071-0	VAR	DANIEL NAVES 3133 SN JOAQUIN	LA SERENA	120			
ABALLAY LUIS ALBA DA CA				A.E. CONTRA EL RIESGO				

- ▶ Muestra: $s \subseteq U$
- ▶ Diseño muestral: $p(s)$ con $s \in \mathcal{S}$ y $\mathcal{S}$ es el espacio de todas las muestras posibles.
- ▶ Probabilidades de inclusión: π_k, π_{kl}
- ▶ Estimador del total de y : $\hat{t}_\pi = \sum_s y_k / \pi_k = \sum_s y_k \times w_k$

No hay marco

- ▶ Construcción del marco $\rightsquigarrow$ Censo
- ▶ Población con características y *tamaño* desconocido
- ▶ Poblaciones pequeñas, ocultas
 - ▶ Explotación sexual comercial de niños, niñas y adolescentes
 - ▶ Niños en situación de calle
 - ▶ Población *trans*
 - ▶ Clasificadores/recicladores, “carritos”

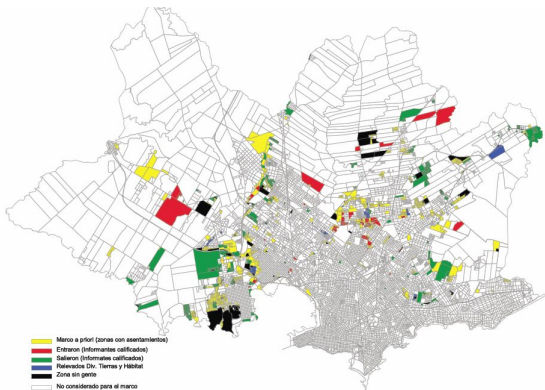


No hay marco: “carritos”

Marcos
imperfectos

Juan José
Goyeneche

- ▶ Usar marco geográfico (si corresponde)
- ▶ Pasos: estratificación, muestra, etc...



Modelos

Conceptos básicos

No existen marcos

Sí existen marcos

Conclusiones

No hay marco: “carritos”

Marcos
imperfectos

Juan José
Goyeneche

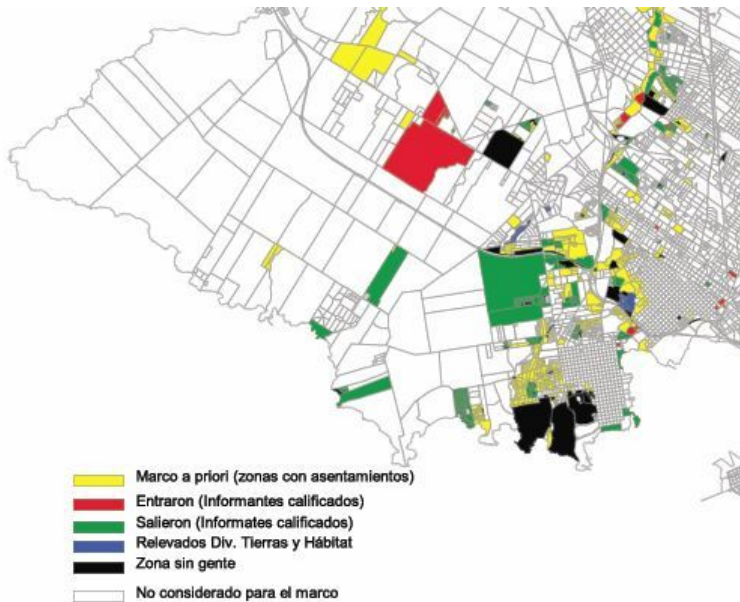
Modelos

Conceptos básicos

No existen marcos

Sí existen marcos

Conclusiones



No hay marco: “trans”

Marcos
imperfectos

Juan José
Goyeneche

Modelos

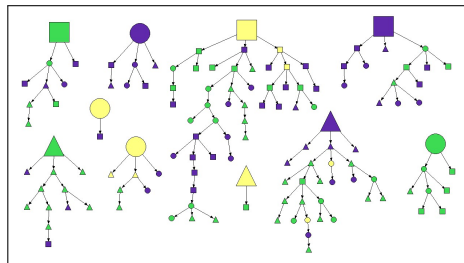
Conceptos básicos

No existen marcos

Sí existen marcos

Conclusiones

- ▶ No usar marco geográfico
- ▶ *Respondent driven sampling* (RDS)



△	MENOR DE 32 AÑOS
○	ENTRE 32 Y 42 AÑOS
□	MAYOR DE 42 AÑOS
■	NO PROSTITUCIÓN
■	PROSTITUCIÓN CALLE
■	PROSTITUCIÓN NO CALLE

Hay marco

Marcos
imperfectos

Juan José
Goyeneche

Modelos

Conceptos básicos

No existen marcos

Sí existen marcos

Conclusiones

- ▶ Registros administrativos
- ▶ Desactualización
- ▶ Sub-cobertura, sobre-cobertura
- ▶ Registros duplicados, “multiplicity”
- ▶ Ejemplos
 - ▶ Padrón electoral
 - ▶ Industrias consumidoras de energía eléctrica
 - ▶ Plan Ceibal: Escuelas, alumnos, maestros, familias.

- ▶ Hay 6.969 circuitos y 2.620.790 habilitados

Modelos

Conceptos básicos

No existen marcos

Sí existen marcos

Conclusiones

 Corte Electoral República Oriental del Uruguay www.corteelectoral.gub.uy					ELECCIONES
PLANES CIRCUITALES					
TACUAREMBO					
Circuito	Urbano / Rural	Serie	Desde	Hasta	Localidad, Local y Dirección
1	Urbano	TAA	1	17568	CIUDAD, IMT-Archivo y Registro de Contribuyentes
2	Urbano	TAA	17569	21073	CIUDAD, IMT-Oficina de Personal
3	Urbano	TAA	21074	23385	CIUDAD, IMT-Oficina de Auditoría
4	Urbano	TAA	23386	25386	CIUDAD, IMT-Oficina Servicios Municipales
5	Urbano	TAA	25387	27148	CIUDAD, Defensoría Pública-Gral. Flores y J.Suarez
6	Urbano	TAA	27149	28716	CIUDAD, Servicio Social-Gral.Artigas 260
7	Urbano	TAA	28717	30077	CIUDAD, Servicio Social-Gral.Artigas 260
8	Urbano	TAA	30078	31442	CIUDAD, BPS-Gral. Artigas 266
9	Urbano	TAA	31443	32854	CIUDAD, BPS-Gral. Artigas 266
10	Urbano	TAA	32855	33955	CIUDAD, Dpto.Mpal. De Cultura
11	Urbano	TAA	33956	35176	CIUDAD, Dpto.Mpal. De Cultura
12	Urbano	TAA	35177	36415	CIUDAD, Dpto.Mpal. De Cultura

Hay marco: Plan circuital

- ▶ Hay 6.969 circuitos y 2.620.790 habilitados



Hay marco: Industrias, electricidad

Marcos
imperfectos

Juan José
Goyeneche

- ▶ Clientes *No Residenciales* de U.T.E (contadores)
- ▶ Más de una empresa por contador
- ▶ Más de un contador por empresa
- ▶ Debo estimar N , recalcular π_k , ...

Modelos

Conceptos básicos

No existen marcos

Sí existen marcos

Conclusiones



Hay marco: Plan Ceibal

Marcos
imperfectos

Juan José
Goyeneche

- ▶ Escuelas (con más de 20 alumnos)
- ▶ Directores
- ▶ Alumnos y grupos (matrícula del año anterior)
- ▶ Maestros: multigrado, multiempleo
- ▶ Familias: $\pi_k \propto$ número de hijos en la escuela

Modelos

Conceptos básicos

No existen marcos

Sí existen marcos

Conclusiones



Marcos imperfectos: conclusiones

Marcos
imperfectos

Juan José
Goyeneche

Modelos

Conceptos básicos

No existen marcos

Sí existen marcos

Conclusiones

- ▶ ¿Debemos “expandir” las correcciones encontradas en la muestra?
- ▶ Si lo hacemos aleatorizamos N_h , N , π_k
- ▶ Como se alteran las fórmulas “básicas”: estimadores de totales, varianzas, etc.
- ▶ La teoría $\neq$ la práctica.
 - ▶ “La realidad supera a ...”
 - ▶ Realidad compleja $\rightsquigarrow$ teoría compleja