

Algoritmos de muestreo y coordinación de muestras

Entrenamiento al INE de Chile

Andrés Gutiérrez - CEPAL

2022

Definiciones básicas

Población finita

Una población finita es un conjunto de N elementos $\{e_1, e_2, \dots, e_N\}$. Cada unidad puede ser identificada sin ambigüedad por un conjunto de rótulos.

Sea $U = \{1, 2, \dots, N\}$ el conjunto de rótulos de la población finita. El tamaño de la población no es necesariamente conocido.

Muestra aleatoria

Es un subconjunto de la población que ha sido extraído mediante un mecanismo estadístico de selección, siguiendo una función de probabilidad. Sin ambigüedad, una muestra¹ seleccionada (realizada) es el conjunto de unidades pertenecientes a

$$s = \{1, \dots, k, \dots, n(S)\}.$$

El número de componentes de s es llamado el *tamaño de muestra* y no siempre es fijo².

¹A mí me gusta más la definición vectorial de una muestra aleatoria

²Es decir, en algunos casos $n(S)$ es una cantidad aleatoria.

Muestra aleatoria sin reemplazo

Una muestra sin reemplazo se denota mediante un vector columna

$$\mathbf{s} = (I_1, I_2, \dots, I_N)' \in \{0, 1\}^N$$

donde

$$I_k = \begin{cases} 1 & \text{si el } k\text{-ésimo elemento pertenece a la muestra,} \\ 0 & \text{en otro caso} \end{cases}$$

Una muestra aleatoria se dice sin reemplazo si la inclusión de cada uno de los elementos se hace entre los elementos que no han sido escogidos aún; de esta manera el conjunto \mathbf{s} nunca tendrá elementos repetidos.

Tamaño de muestra

El tamaño de muestra corresponde a la cardinalidad de s .

$$n(S) = \sum_{k \in U} I_k.$$

Como $n(S)$ no es una cantidad fija, es posible que ocurran uno de los siguientes escenarios:

- a. que la muestra no contenga a ningún elemento, entonces esta muestra se dice vacía;
- b. que la muestra contenga a todos los elementos de la población, esta muestra se conoce con el nombre de *censo*.

Soporte

El conjunto de todas las posibles muestras se conoce como *soporte* (espacio muestral).

- ▶ Un *soporte* Q es un conjunto de muestras.
- ▶ Un soporte se llama *simétrico* si para cualquier $s \in Q$, todas las permutaciones de s están también en Q .

Dos soportes importantes

El *soporte simétrico sin reemplazo* se define como

$$\mathcal{S} = \{0, 1\}^N$$

Nótese que

$$\#(\mathcal{S}) = 2^N$$

Por ejemplo, si $N = 3$, entonces \mathcal{S} queda definido por las siguientes muestras:

$$\begin{aligned}\mathcal{S} = \{ & (0, 0, 0)', \\ & (1, 0, 0)', (0, 1, 0)', (0, 0, 1)', \\ & (1, 1, 0)', (1, 0, 1)', (0, 1, 1)', \\ & (1, 1, 1)' \}\end{aligned}$$

Dos soportes importantes

El *soporte simétrico sin reemplazo de tamaño fijo* definido como

$$\mathcal{S}_n = \left\{ \mathbf{s} \in \mathcal{S} \mid \sum_{k \in U} s_k = n \right\}$$

Nótese que

$$\#(\mathcal{S}_n) = \binom{N}{n}$$

Por ejemplo, si $N = 3$ y $n = 2$, entonces \mathcal{S}_n queda definido por las siguientes muestras:

$$\mathcal{S}_n = \{(1, 0, 1)', (1, 1, 0)', (0, 1, 1)'\}$$

Diseño de muestreo

Un *diseño de muestreo* $p(\cdot)$ es una distribución de probabilidad multivariante definida sobre un soporte Q ; es decir, $p(\cdot)$ es una función que va desde Q hasta $(0, 1]$ tal que $p(s) > 0$ para todo $s \in Q$ y

$$\sum_{s \in Q} p(s) = 1$$

Otra vez:

Dado el soporte Q , un diseño de muestreo es una función $p(\cdot)$, tal que $p(s)$ arroja la probabilidad de selección de la muestra realizada s bajo un esquema de selección particular. En otras palabras, si S es una muestra aleatoria que toma el valor s , entonces

$$Pr(S = s) = p(s) \quad \text{para todo } s \in Q.$$

Por consiguiente, $p(\cdot)$ es llamada diseño de muestreo.

Algoritmo y diseño

- ▶ Nótese que el diseño de muestreo no se refiere a un algoritmo o procedimiento que permita la selección de muestras.
- ▶ Dado un diseño de muestreo, el trabajo del estadístico consiste en encontrar un algoritmo que permita la selección de muestras cuya probabilidad de selección corresponda a la probabilidad inducida por el diseño de muestreo.

Algoritmos de selección

Un algoritmo de selección es un procedimiento usado para seleccionar una muestra probabilística.

Es imposible listar todas las posibles muestras, generar una variable aleatoria con distribución uniforme en el intervalo $[0, 1]$ para luego hacer la correspondiente selección.

Variable indicadora

La inclusión del elemento k -ésimo en una muestra s particular es un evento aleatorio definido por la función indicadora $I_k(s)$, que está dada por

$$I_k(s) = \begin{cases} 1 & \text{si } k \in s \\ 0 & \text{si } k \notin s. \end{cases}$$

Probabilidad de inclusión

Bajo un diseño de muestreo $p(\cdot)$, para el elemento k -ésimo de la población, la probabilidad de inclusión se denota como π_k y se conoce como la probabilidad de inclusión de **primer orden** y está dada por

$$\pi_k = Pr(k \in S) = Pr(I_k = 1) = E_p(I_k = 1) = \sum_{s \ni k} p(s). \quad (1)$$

Algunos diseños y algoritmos de muestreo

Diseño de muestreo aleatorio simple (MAS)

Un diseño de muestreo se dice aleatorio simple sin reemplazo si todas las posibles muestras de tamaño n tienen la misma probabilidad de ser seleccionadas. Así,

$$p(s) = \begin{cases} \frac{1}{\binom{N}{n}} & \text{si } \#(s) = n \\ 0 & \text{en otro caso} \end{cases} \quad (2)$$

Nótese que $\sum_{s \in Q} p(s) = 1$ porque $\#(Q) = \binom{N}{n}$.

Algoritmo secuencial de selección y rechazo (FMR - 1962)

Algorithm 4.3 Selection-rejection procedure for SRSWOR

DEFINITION k, j : INTEGER;

$j = 0$;

FOR $k = 1, \dots, N$ DO

with probability $\frac{n-j}{N-(k-1)}$ THEN

select unit k ;
 $j = j + 1$;

ENDFOR.

Ejemplo

Considere un población de tamaño $N = 5$ y un diseño de muestreo MAS de tamaño de muestra $n = 2$

k	ξ_k^P	ck	j	s
1	0.4938	?	?	?
2	0.7044	?	?	?
3	0.4585	?	?	?
4	0.6747	?	?	?
5	0.8565	?	?	?

Complete las celdas faltantes y seleccione la muestra.

Ejemplo

Considere un población de tamaño $N = 5$ y un diseño de muestreo MAS de tamaño de muestra $n = 2$

k	ξ_k^P	ck	j	s
1	0.4938	0.4000000	0	0
2	0.7044	0.5000000	0	0
3	0.4585	0.6666667	1	1
4	0.6747	0.5000000	1	0
5	0.8565	1.0000000	2	1

¿Este algoritmo induce las mismas probabilidades de selección de un MAS?

Verificación simplificada

$$\begin{aligned}Pr(S = s) &= Pr(S = (0, 0, 1, 0, 1)') \\&= Pr(I_1 = 0, I_2 = 0, I_3 = 1, I_4 = 0, I_5 = 1) \\&= Pr(I_5 = 1 | I_1 = 0, I_2 = 0, I_3 = 1, I_4 = 0) \\&\times Pr(I_4 = 0 | I_1 = 0, I_2 = 0, I_3) \\&\times Pr(I_3 = 1 | I_1 = 0, I_2 = 0) \\&\times Pr(I_2 = 0 | I_1 = 0) \\&\times Pr(I_1 = 0)\end{aligned}$$

Verificación simplificada

$$\begin{aligned}Pr(S = s) &= Pr\left(u < \frac{2-1}{5-(5-1)}\right) Pr\left(u > \frac{2-1}{5-(4-1)}\right) \\&\quad Pr\left(u < \frac{2-0}{5-(3-1)}\right) Pr\left(u > \frac{2-0}{5-(2-1)}\right) \\&\quad Pr\left(u > \frac{2-0}{5-(1-1)}\right) \\&= \frac{1}{1} \times \frac{1}{2} \times \frac{2}{3} \times \frac{1}{2} \times \frac{3}{5} \\&= \frac{1}{10} = \frac{1}{\binom{5}{2}} = \binom{N}{n}^{-1}\end{aligned}$$

Demostración

Proof: From Algorithm 2, we directly obtain

$$\begin{aligned} \Pr(S = s) &= \left[\prod_{k \in s} \Pr(k \in s | n_{k-1} = \#\{j \in s \mid j < k\}) \right] \\ &\quad \times \prod_{k \in U \setminus s} \Pr(k \notin s | n_{k-1} = \#\{j \in s \mid j < k\}) \\ &= \left[\prod_{k \in s} \frac{n - \#\{j \in s \mid j < k\}}{N - (k - 1)} \right] \times \prod_{k \in U \setminus s} \left[1 - \frac{n - \#\{j \in s \mid j < k\}}{N - (k - 1)} \right] \\ &= \left[\prod_{k \in U} (N - k + 1) \right]^{-1} \times \left[\prod_{k \in s} (n - \#\{j \in s \mid j < k\}) \right] \\ &\quad \times \left[\prod_{k \in U \setminus s} (N - k + 1 - n + \#\{j \in s \mid j < k\}) \right] \\ &= \frac{n!(N - n)!}{N!} = \binom{N}{n}^{-1}. \end{aligned}$$

Algoritmo de ordenamiento aleatorio (Sunter - 1977)

Algorithm 4.5 Random sort procedure for SRSWOR

1. A value of an independent uniform variable in $[0,1]$ is allocated to each unit of the population.
 2. The population is sorted in ascending (or descending) order.
 3. The first (or last) n units of the sorted population are selected in the sample.
-

Demostración

Result 13. *The random sorting method gives a SRSWOR.*

Proof. Consider the distribution function of the largest uniform random number generated for $N - n$ units. If u_i denotes the uniform random variable of the i th unit of the file, this distribution function is given by:

$$F(x) = \Pr \left[\bigcap_{i=1}^{N-n} (u_i \leq x) \right] = \prod_{i=1}^{N-n} \Pr(u_i \leq x) = x^{N-n}, \quad 0 < x < 1.$$

The probability that the n remaining units are all greater than x is equal to $(1 - x)^n$. The probability of selecting a particular sample \mathbf{s} is thus

$$p(\mathbf{s}) = \int_0^1 (1 - x)^n dF(x) = (N - n) \int_0^1 (1 - x)^n x^{N-n-1} dx, \quad (4.6)$$

which is an Euler integral

$$B(x, y) = \int_0^1 t^x (1 - t)^y dt = \frac{x!y!}{(x + y + 1)!}, \quad x, y \text{ integer.} \quad (4.7)$$

From (4.6) and (4.7), we get

$$p(\mathbf{s}) = \binom{N}{n}^{-1}. \quad \square$$

Coordinación de muestras

Elementos básicos

Considere una población finita investigada en dos ocasiones, denotadas como U^1 y U^2 en los respectivos periodos. En cada ocasión, se selecciona una muestra s^t de U^t , $t = 1, 2$.

- ▶ Una muestra es *positivamente coordinada* con otra, si comparten todos sus elementos.
- ▶ Dos muestras son *negativamente coordinadas* si no comparten ningún elemento en común.

Coordinación positiva

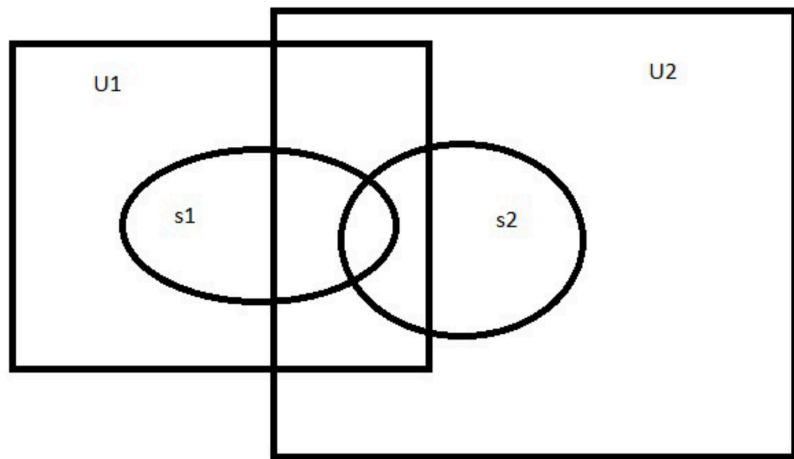


Figure 1: Matei & Grafström (2019)

Coordinación negativa

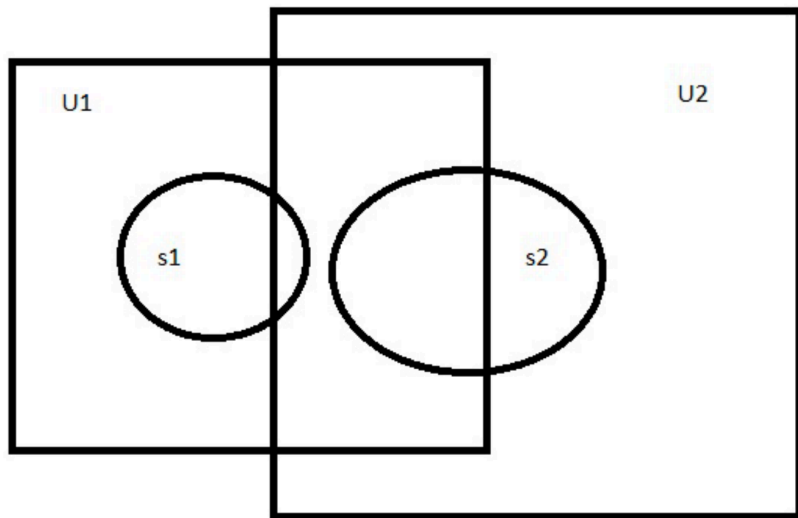


Figure 2: Matei & Grafström (2019)

Tipos de coordinación

Existen básicamente dos tipos de coordinación: la primera basada en números aleatorios permanentes (NAP) y la segunda basada en programación (lineal o no).

Tipos de coordinación

- ▶ NAP: a cada unidad de la población se le asigna un número aleatorio uniforme que será utilizado para efectuar todos los procesos de selección durante un periodo de tiempo definido. Ernst (1999), Mach et al. (2006), Nedyalkova et al.(2008), Nedyalkova et al.(2009), Grafström and Matei (2015), Grafström and Matei (2018).
- ▶ Programación matemática y otros métodos: Keyfitz (1951), Raj (1968), Kish and Scott (1971), Arthnari and Dodge (1981), Causey et al. (1985), Ernst and Ikeda (1995), Ernst (1996, 1998), Ernst and Paben (2002), Matei and Tillé (2005), Mach et al. (2006), Matei and Skinner (2009), Schiopu-Kratina et al. (2014).

Números aleatorios permanentes

Cada unidad del marco recibirá un número aleatorio con distribución uniforme en el intervalo unitario. Es decir a cada unidad $k \in U$ se le asignara el siguiente número:

$$\xi_k^P \sim \textit{Uniforme} (0, 1)$$

Evidentemente, en este caso los números aleatorios permanentes no son equidistantes.

Números aleatorios colocados

A partir de los números aleatorios ξ_k creados en el paso anterior, es posible utilizar su rango para definir su ordenamiento y mediante la siguiente función crear números aleatorios equidistantes:

$$\xi_k^C = \frac{\text{Rango}(\xi_k^P) - \varepsilon}{N}$$

En donde ε es un único valor aleatorio entre cero y uno.

Ejemplo

Como ilustración, considere una población de tamaño $N = 10$, para la que se han definido números aleatorios ξ_i^P y, teniendo en cuenta un número aleatorio $\varepsilon = 0.283$ también se definen ξ_i^C .

unit	xi_P	xi_C
1	0.2875	0.1717
2	0.7883	0.6717
3	0.4089	0.2717
4	0.8831	0.7717
5	0.9404	0.9717
6	0.0455	0.0717
7	0.5281	0.4717
8	0.8924	0.8717
9	0.5514	0.5717
10	0.4566	0.3717

Coordinación de muestras aleatorias simples

Algorithm 4.1 Bernoulli sampling without replacement

DEFINITION k : INTEGER;

FOR $k = 1, \dots, N$ DO with probability π select unit k ; ENDFOR.

BER

```
N <- 10
n <- 3
pik <- n / N

# Random sample size
table1 %>%
  select(-xi_C) %>%
  mutate(s1 =
    case_when(xi_P > 0 & xi_P <= pik ~ 1,
              TRUE ~ 0))
```

BER

unit	ξ_P	s1
1	0.2875	1
2	0.7883	0
3	0.4089	0
4	0.8831	0
5	0.9404	0
6	0.0455	1
7	0.5281	0
8	0.8924	0
9	0.5514	0
10	0.4566	0

Coordinación

Para coordinar dos muestras s^1 de tamaño n_1 y s^2 de tamaño n_2 , es posible:

1. Escoger dos constantes a_1 y a_2 en el intervalo $(0, 1)$.
2. A partir del marco ordenado con los números aleatorios permanentes (o colocados), definir la muestra s_1 como las primeras n_1 unidades a la derecha (o izquierda) de a_1 y la muestra s_2 como las primeras n_2 unidades a la derecha (o izquierda) de a_2 .

Coordinación

- ▶ Si se quieren muestras positivamente coordinadas, entonces $a_1 = a_2$.
- ▶ Por el contrario, si se quieren muestras negativamente coordinadas, se deberán escoger las constantes de forma apropiada.
 - ▶ Por ejemplo sumar 0.5 (en módulo uno) a la constante a_1 ; es decir, $a_2 = (a_1 + 1/2) \bmod 1$.
 - ▶ En general, si se quieren coordinar negativamente Q diferentes muestras, @Grafstrom_Matei_2015 aconsejan añadir la cantidad de $1/Q$ (en módulo uno) a la constante a_1 .

BER

El tamaño de muestra de este diseño es aleatorio.

```
n <- 3
Q <- 2
a1 <- 0
a2 <- a1 + 1 / Q

tableBER <- table1 %>%
  select(-xi_C) %>%
  mutate(
    s1 = case_when(xi_P > a1 & xi_P <= a1 + n / N ~ 1,
                   TRUE ~ 0),
    s2 = case_when(xi_P > a2 & xi_P <= a2 + n / N ~ 1,
                   TRUE ~ 0)
  )
```

BER

unit	xi_P	s1	s2
1	0.2875	1	0
2	0.7883	0	1
3	0.4089	0	0
4	0.8831	0	0
5	0.9404	0	0
6	0.0455	1	0
7	0.5281	0	1
8	0.8924	0	0
9	0.5514	0	1
10	0.4566	0	0

MAS con NAC

El tamaño de muestra de este diseño es fijo.

```
tableSRS1 <- table1 %>%  
  select(-xi_P) %>%  
  mutate(  
    s1 = case_when(xi_C > a1 & xi_C <= a1 + n / N ~ 1,  
                   TRUE ~ 0),  
    s2 = case_when(xi_C > a2 & xi_C <= a2 + n / N ~ 1,  
                   TRUE ~ 0)  
  )
```

MAS con NAP

unit	xi_C	s1	s2
1	0.1717	1	0
2	0.6717	0	1
3	0.2717	1	0
4	0.7717	0	1
5	0.9717	0	0
6	0.0717	1	0
7	0.4717	0	0
8	0.8717	0	0
9	0.5717	0	1
10	0.3717	0	0

MAS con NAP

El tamaño de muestra de este diseño es fijo

```
tableSRS2 <- table1 %>%  
  select(-xi_C) %>%  
  arrange(xi_P) %>%  
  mutate(s1 = c(rep(1, 3), rep(0, 7)),  
         xi_P2 = (xi_P + a2) %% 1) %>%  
  arrange(xi_P2) %>%  
  mutate(s2 = c(rep(1, 3), rep(0, 7))) %>%  
  arrange(unit)
```

MAS con NAP

unit	xi_P	s1	xi_P2	s2
1	0.2875	1	0.7875	0
2	0.7883	0	0.2883	1
3	0.4089	1	0.9089	0
4	0.8831	0	0.3831	0
5	0.9404	0	0.4404	0
6	0.0455	1	0.5455	0
7	0.5281	0	0.0281	1
8	0.8924	0	0.3924	0
9	0.5514	0	0.0514	1
10	0.4566	0	0.9566	0

Coordinación de muestras proporcionales a una
medida de tamaño

Algorithm 5.4 Sequential procedure for POISSWOR

DEFINITION k : INTEGER;

FOR $k = 1, \dots, N$, DO select the k th unit with probability π_k ;
ENDFOR.

En donde π_k es la probabilidad de inclusión calculada con base en la información auxiliar del marco (medida de tamaño):

$$\pi_k = n \frac{x_k}{\sum_U x_k}$$

Equivalencia

En el algoritmo anterior, $k \in s$ sí y solo sí

$$\xi_k^P \leq n \frac{x_k}{\sum_U x_k} = \pi_k$$

Teniendo en cuenta que:

$$\xi_k^P \leq \pi_k \Leftrightarrow \frac{\xi_k^P}{N \times p_k} \leq \frac{\pi_k}{N \times p_k}$$

Entonces, equivalentemente $k \in s$ sí y solo sí

$$\xi_i^{pps} \leq \frac{n}{N}$$

En donde $p_k = \pi_k/n$ y $\xi_i^{pps} = \frac{\xi_k^P}{N \times p_k}$.

```
table1 %>%  
  select(-xi_C) %>%  
  mutate(  
    xk = c(198, 173, 184, 179, 170,  
           190, 162, 159, 166, 190),  
    pik = n * xk / sum(xk),  
    xi_pps = round(xi_P / (N * pik / n), 4),  
    s1 = case_when(xi_pps > 0 & xi_pps <= (n / N) ~ 1,  
                   TRUE ~ 0)  
  )
```

unit	xi_P	xk	pik	xi_pps	s1
1	0.2875	198	0.3354037	0.2572	1
2	0.7883	173	0.2930548	0.8070	0
3	0.4089	184	0.3116883	0.3936	0
4	0.8831	179	0.3032185	0.8737	0
5	0.9404	170	0.2879729	0.9797	0
6	0.0455	190	0.3218521	0.0424	1
7	0.5281	162	0.2744212	0.5773	0
8	0.8924	159	0.2693394	0.9940	0
9	0.5514	166	0.2811971	0.5883	0
10	0.4566	190	0.3218521	0.4256	0

Coordinación

- ▶ Al ordenar el marco mediante los números ξ_i^{pps} y seleccionar los primeros elementos, se obtendrá una muestra secuencial de Poisson.
- ▶ Es posible aplicar los mismos principios de la sección anterior. Es decir, para coordinar dos muestras s^1 de tamaño n_1 y s^2 de tamaño n_2 , es posible escoger dos constantes a_1 y a_2 en el intervalo $(0, 1)$.
- ▶ Luego, a partir del marco ordenado, definir la muestra s_1 como las primeras n_1 unidades a la derecha (o izquierda) de a_1 y la muestra s_2 como las primeras n_2 unidades a la derecha (o izquierda) de a_2 .

SPO

```
tableSP01 <- table1 %>%  
  mutate(  
    xk = c(198, 173, 184, 179, 170,  
           190, 162, 159, 166, 190),  
    pik = n * xk / sum(xk),  
    xi_pps = round(xi_P / (N * pik / n), 4)  
  ) %>%  
  arrange(xi_pps) %>%  
  mutate(s1 = c(rep(1, 3), rep(0, 7)),  
         xi_pps2 = (xi_pps + a2) %% 1) %>%  
  arrange(xi_pps2) %>%  
  mutate(s2 = c(rep(1, 3), rep(0, 7))) %>%  
  arrange(unit) %>%  
  select(-c(xi_C, xk, xi_P))
```

SPO

unit	pik	xi_pps	s1	xi_pps2	s2
1	0.3354037	0.2572	1	0.7572	0
2	0.2930548	0.8070	0	0.3070	1
3	0.3116883	0.3936	1	0.8936	0
4	0.3032185	0.8737	0	0.3737	0
5	0.2879729	0.9797	0	0.4797	0
6	0.3218521	0.0424	1	0.5424	0
7	0.2744212	0.5773	0	0.0773	1
8	0.2693394	0.9940	0	0.4940	0
9	0.2811971	0.5883	0	0.0883	1
10	0.3218521	0.4256	0	0.9256	0

PAR

En Brasil, el IBGE utiliza el algoritmo de Pareto para la selección de muestras coordinadas en la PNADC. Este algoritmo hace uso de los principios de la función de distribución de Pareto con parámetros $(1, 1)$ y crea los siguientes números aleatorios permanentes:

$$\xi_k^{par} = \frac{\xi_k^P / (1 - \xi_k^P)}{\pi_k / (1 - \pi_k)}$$

En donde $\pi_k = n \frac{x_k}{\sum_U x_k}$.



On sampling with probability proportional to size

Bengt Rosén*

Statistics Sweden, S-115 81 Stockholm, Sweden

Received 7 June 1996

Abstract

One of the vehicles for utilization of auxiliary information is to use a sampling scheme with inclusion probabilities proportional to given size measures, a π ps scheme. The paper addresses the following π ps problem: Exhibit a π ps scheme with prescribed sample size, which leads to good estimation precision and has good variance estimation properties.

Rosén (1997) presented a novel general class of sampling schemes, called order sampling schemes, which here are shown to provide interesting contributions to the π ps problem. A notion 'order sampling with fixed distribution shape' (OSFS) is introduced, and employed to construct a general class of π ps schemes, called OSFS π ps schemes. A particular scheme, Pareto π ps, is shown to be optimal among OSFS π ps schemes, in the sense that it minimizes estimator variances. Comparisons are made of three OSFS π ps schemes and three other π ps schemes; Sunter π ps and systematic π ps with frame ordered at random respectively by the sizes. The main conclusion is as follows. Pareto π ps is superior among π ps schemes which admit objective assessment of sampling errors. © 1997 Elsevier Science B.V.

PAR

```
tablePA <- table1 %>%  
  mutate(  
    xk = c(198, 173, 184, 179, 170,  
           190, 162, 159, 166, 190),  
    pik = n * xk / sum(xk),  
    xi_par = round((xi_P * (1 - pik)) / (pik * (1 - xi_P)))  
  ) %>%  
  arrange(xi_par) %>%  
  mutate(s1 = c(rep(1, 3), rep(0, 7))) %>%  
  arrange(s1, xi_par) %>%  
  mutate(s2 = c(rep(1, 3), rep(0, 7))) %>%  
  arrange(unit) %>%  
  select(-c(xi_C, xk, xi_P))
```

PAR

unit	pik	xi_par	s1	s2
1	0.3354037	0.7995	1	0
2	0.2930548	8.9827	0	0
3	0.3116883	1.5276	1	0
4	0.3032185	17.3595	0	0
5	0.2879729	39.0132	0	0
6	0.3218521	0.1004	1	0
7	0.2744212	2.9589	0	1
8	0.2693394	22.4990	0	0
9	0.2811971	3.1420	0	1
10	0.3218521	1.7705	0	1

¿Qué sigue?

Para la próxima sesión

1. Encuestas de hogares
 - ▶ Grupos rotacionales
2. Encuestas de establecimientos
 - ▶ Nacimientos
 - ▶ Defunciones
 - ▶ Cambios en la clasificación
3. Algoritmos en R

Muchas gracias

¡Gracias!

Email: andres.gutierrez@cepal.org