

##### HHIS I

2000. július 16.

**8. verzió**

## Tartalom

*Mintavételi terv*

oldal

Rövid összefoglalás 3

A mintavétel módszerének, folyamatának részletes ismertetése 4

#### Mintavételi terv

**Rövid összefoglalás**

* **A vizsgálati populáció:**

Magyarország 1999. december 31-ig 18. életévét betöltött lakossága. **N=8 180 439**

* **A minta nagysága**:n**=** 7000 fő.
* **A mintavétel típusa**: kétlépcsős csoportos mintavétel

**1.** **lépcső**

**700** db 10 fős csoport és ezen keresztül *települések* kiválasztása **lakosságszámmal arányos** **szisztematikus visszatevés nélküli mintavétellel.**

*A módszer rövid ismertetése*:

* + 1. Szisztematikus mintavétellel kiválasztjuk a nagy létszámú célpopulációt tartalmazó településeket, amelyek biztosan bekerülnek a mintába.
    2. A megmaradt települések közül véletlen szisztematikus mintavétellel kiválasztjuk azokat, amelyek még a mintába kerülhetnek.

**Eredménye**: ***Településminta* =** *mintatelepülések listája és az egyes mintatelepülésekről a második lépcsőben kiválasztandó személyek száma.*

A mintatelepülések az *elsődleges mintavételi egységek*.

**2. lépcső**

*Minta-településenként* egyszerű visszatevés nélküli mintavétellel kiválasztjuk - az első lépcsőben meghatározott számú - mintába kerülő személyeket.

**Eredménye:** Minta-településenkénti ***címlista*.**

* **A mintavételhez szükséges országos adatbázisok**

1. lépcső:

A települések kiválasztásához szükséges **mintakeret:**

* 1. a vizsgálati populáció *településeinek listája* (település kóddal), 3153 település
  2. a települések 1999. december 31-ig 18. életévét betöltött lakosainak száma, “**nem” (férfi-nő)** szerinti bontásban (a születési év nem lehet nagyobb 1981-nél)
  3. a települések megyéje (a megye kódja)

*Forrás*: Népességnyilvántartó Hivatal (továbbiakban NH)

Cím: Dr. Szalay László főosztályvezető, BM Közp. Hivatal, Igazgatási és Felügyeleti Főosztály , 1450 Budapest, Pf. 81 (Balázs B. u.)

Tel.: 455-6781, Mobil: 20/9846670

Fax: 455-6779

“Contact person”: Dr. Cziffra András, *mobil*: 20/955-4137, *e-mail*:

[andras.cziffra@konyv.kozig.b-m.hu](mailto:andras.cziffra@konyv.kozig.b-m.hu) (közvetlen: Varga János informatikus, 455 6700)

*Kapcsolódó EFKI Eg. Stat. Egys.* (továbbiakban EFKI ESE) a*datbázis neve:* **tminta**

(ld. Országh Sándornál)

2. lépcső

A minta-települések 1999. dec. 31-ig 18. életévét betöltött lakosainak neve, címe, neme és születési éve

*Forrás*: Népességnyilvántartó Hivatal

Cím: Dr. Szalay László főosztályvezető, BM Közp. Hivatal, Igazgatási és Felügyeleti Főosztály, 1450 Budapest, Pf. 81 (Balázs B. u.)

Tel.: 455-6781, Mobil: 20/9846670

Fax: 455-6779

„Contact person”: Dr. Cziffra András, *mobil*: 20/955-4137, *e-mail*: [andras.cziffra@konyv.kozig.b-m.hu](mailto:andras.cziffra@konyv.kozig.b-m.hu) (közvetlen: Varga János informatikus, 455 6700)

*Kapcsolódó EFKI ESE adatbázis neve:* **clista** (ld. Országh Sándornál)

* **Az adatok áramlásának folyamata**

# Mintakeret

*NH→ EFKI ESE*

# Településminta (*tminta)*

*EFKI ESE → NH*

***Címlista***

NH → EFKI ESE ***(clista)***

###### A mintavétel módszerének, folyamatának részletes ismertetése

**1. Lépcső:** a települések(10 fős csoportok)lakosságszámmal arányos, szisztematikus, visszatevés nélküli mintavétellel történő kiválasztása.

1..1. Az első fázisban *nem –véletlen mintavételt* alkalmazunk a nagy települések *biztos kiválasztására*.

1.1.1. Először kiszámoljuk a kiválasztási intervallumot:

,

*I*-ben a kitevő helyén az (1)-es az első ciklust jelöli.

**Eredmény:** Megkapjuk azokat a településeket, amelyek célpopulációjának nagysága meghaladja a kiválasztási intervallum értékét. Legyen ezek száma az első ciklusban  (a kitevő helyén az (1)-es az első ciklust jelöli).

1.1.2. Minden biztosan kiválasztott település célpopulációjának nagyságát (elosztjuk a kiválasztási intervallummal :



**Eredmény:** Megkapjuk, hogy minden egyes biztosan kiválasztott településről hány db 10 fős csoportot válasszunk ki:



Ha nem egész számot kapunk, akkor a matematikai szabályoknak

megfelelően, a legközelebbi egész számra kerekítünk.

1.2. A biztosan kiválasztott településeken a kapott 10 fős csoportok számát megszorozzuk 10-zel, s megkapjuk, hogy hány személyt kell kiválasztani a településről ().

1. 3. A biztosan kiválasztott településeket töröljük a listáról és az összegezett célpopulációjuk nagyságával csökkentjük a célpopuláció nagyságát. A redukált célpopulációval és a redukált település-számmal, illetve a redukált 10 fős csoportok számával megismételjük az 1.1. fázisban ismertetett eljárást.

Akkor fejezzük be az ismétlést, ha már nem találunk olyan településeket, amelyek célpopulációja nagyobb, mint a kiválasztási intervallum.

**Az első fázis (nem-véletlen mintavétel) eredménye:**

* A biztosan kiválasztott települések mintája (lista + a mintatelepülésről kiválasztandó személyek száma)
* A redukált célpopuláció = (a célpopuláció)⎯(a biztosan kiválasztott települések célpopulációjának összege)≡N-Lb=8 180 439-Lb
* A 10 fős csoportok redukált száma = 700⎯(az összes biztosan kiválasztott településre eső 10 fős csoportok száma.)≡700-[[1]](#footnote-1)☝

1. Tábla: *Az első fázis eredményének összefoglalója*

|  |  |  |  |
| --- | --- | --- | --- |
| A biztosan kiválasztott település neve | A biztosan kiválasztott település célpopulációjának nagysága | A biztosan kiválasztott településekbe eső 10 fős csoportok száma | A biztosan kiválasztott településekről a mintába kerülők száma |
|  | L1b |  | \*10 |
|  | L2b |  | \*10 |
|  |  |  |  |
| Összesen | Lb |  | \*10 |

2.1. A második fázisban *véletlen szisztematikus mintavételt* alkalmazunk a mintába még bekerülhető települések kiválasztására.

* Az országos szintű redukált célpopulációt tekintjük: Lv=N-Lb=8 180 439-Lb
* A megyéket véletlenszerűen sorba rendezzük.
* Minden megyén belül településnagyság-kategóriákat hozunk létre:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *A települések célpopuláció nagyság szerinti kategóriái* | | | | | |
| 0-1000 | 1001-3000 | 3001-5000 | 5001-10000 | 10001-50000 | 50001-X |

* A településnagyság–kategóriákon belül a megmaradt településeket *véletlen sorrendbe* rendezzük.
* Kiszámítjuk a kiválasztási intervallumot:



* 1 és *Iv*között véletlen számot választunk. Ez egy lakosságszámot jelent, amellyel kijelöljük az első kiválasztandó települést. Jelöljük *Vsz*-szel.
* A Vsz, Vsz+Iv, Vsz+2Iv, … lépésekkel kapott lakosságszámok kijelölik a kiválasztandó településeket és azt, hogy egy kiválasztott településről hány 10 fős csoportot választhatunk (egy településből többször is választhatunk 10 fős csoportot, ha a lakosság száma legalább kétszer akkora, mint a kiválasztási intervallum)

**A második fázis (véletlen mintavétel) eredménye**

A véletlenül kiválasztott települések mintája (lista + kiválasztandó személyek száma)

2. Tábla: *A második fázis eredményének összefoglalója*

|  |  |  |  |
| --- | --- | --- | --- |
| A véletlenül kiválasztott település neve | A véletlenül kiválasztott település célpopulációjának nagysága | A véletlenül kiválasztott településekbe eső 10 fős csoportok száma | A véletlenül kiválasztott településekről a mintába kerülők száma |
|  | L1v |  | \*10 |
|  | L2v |  | \*10 |
|  |  |  |  |
| Összesen | Lv |  | \*10 |

**Az első lépcső eredménye**: a biztosan és a véletlenül kiválasztott településminta együtt= ***településminta***→*elsődleges mintavételi egységek* (PSUs): mintatelepülések

**2. Lépcső**

*Minta-településenként* egyszerű visszatevés nélküli mintavétellel kiválasztjuk – a második lépésben meghatározott számú - mintába kerülő személyeket.

**A második lépcső eredménye:** Minta-településenkénti **címlista.**

**Kiválasztási valószínűségek** *(selection probability)*

* *A biztosan kiválasztott településekről a mintába került személyek kiválasztási valószínűsége*:

1. annak a valószínűsége, hogy a biztosan kiválasztott településre eső *i*-edik csoportot választjuk ki:



1. annak a valószínűsége, hogy egy személy az *i-*edik biztosan kiválasztott településről kerül kiválasztásra (benne lesz egy 10 fős csoportban):



1. a biztosan kiválasztott településekről mintába került személyek kiválasztási valószínűsége:

**

ahol Csb a biztosan kiválasztott települések 10-es csoportjai számának

összege.[[2]](#footnote-2)🙵

* *A véletlenül kiválasztott településekről a mintába került személyek kiválasztási valószínűsége*:

1. annak a valószínűsége, hogy az *i*-edik települést választjuk ki:



1. annak a valószínűsége, hogy egy személy az *i-*edik véletlenül kiválasztott településről kerül kiválasztásra (benne lesz egy 10 fős csoportban):



1. a véletlenül kiválasztott településekről a mintába került személyek kiválasztási valószínűsége:

**

**A mintavételi súlyok** *(sampling* *weight):* a kiválasztási valószínűségek reciproka

**Irodalom**

P S Levy-S Lemeshow: *Sampling of Populations*, John Wiley &Sons, Inc. 1999

1. ☝ Ez kicsit eltérhet a második fázisban kiválasztott 10 fős csoportok számától () a véletlen kiválasztás miatt. [↑](#footnote-ref-1)
2. 🙵 A csoportok végső összegével számolunk. [↑](#footnote-ref-2)