

A mintavételi terv elemei

Célsokaság: legalább 14 éves, nem intézeti népesség.

Mintavételi keret: BM Személyiadat- és lakcímnnyilvántartás (januári állapot)

A tervezett minta valószínűségi minta a mintavételi keretből, ami azt jelenti, hogy minden személynek ismert és pozitív a mintába kerülési valószínűsége. Ez teszi lehetővé, hogy a mintából olyan becslések szülessenek, amiknél használható a matematikai statisztika eszköztára, a többi között a becslések pontosságának jellemzésére.

A tervezett elemszámot az határozza meg, hogy minden régióra azonos pontosságú becslések szülessenek, ezért a régiós várható megvalósult elemszámok egyenlők (1600).

A H-UNCOVER felvétel mintája kétlépcsős, ahol az első lépcsőben településeket választottunk rétegzés mellett nagysággal arányos valószínűséggel, a második lépcsőben a kiválasztott településeken személyeket választottunk véletlen szisztematikus módon.

A településminta

Az ország 3177 településéből (Budapest kerületeit külön számolva) 489-et választottunk ki az alábbi módon.

Budapest kerületei és minden régión belül a nagyobb települések (ahol az 1600-as régiós megvalósult elemszámból legalább 10 jut arányos allokáció mellett, a település méretéből adódóan) 1 valószínűséggel kerültek a mintába, összesen 170. Ugyancsak 1 valószínűséggel választottuk mintába azt a 11 települést, ahol az elérhető információk szerint legalább 5 megbetegedést regisztráltak. Ezzel 181 ún. önreprezentáló településünk lett, ami a teljes célsokaság 55%-át fedi le, illetve az (akkori adatok alapján) az összes regisztrált fertőzött 82%-át. Ezekben a településeken hatékonyabb, egylépcsős kiválasztással jutunk a személyekhez.

A maradék közel 3000 településből 308-at választottunk nagysággal arányos valószínűségű kiválasztással, rétegzés mellett.

Minden régión belül rétegző változó volt annak indikátora, hogy adott településen regisztráltak-e fertőzöttet. Az így létrehozott rétegeket tovább rétegeztük a település nagysága, és ahol a réteg mérete megengedte, az egy főre eső SZJA alapot képező jövedelem és a felsőfokú végzettségűek aránya szerint. Összesen 154 réteget képeztünk, minden rétegből 2 települést választottunk ki nagysággal arányos valószínűséggel.

A személyek kiválasztása

A kiválasztott településeken a személyeket a keretből véletlen szisztematikus módon választjuk ki, a keretet a személyek születési dátuma szerint sorba rendezve. Az életkor és a fertőzöttség közötti jelentősebb összefüggés esetén ez a kiválasztási mód hatékonyabb mintát, pontosabb becsléseket eredményez.

A személyi keretből kizártuk a 110 évnél idősebbeket (kerethibát feltételezve) és olyan címen lakókat, ahol legalább 20 személynek van bejelentett lakóhelye (ezzel az intézeti népesség jelentős részét zártuk ki a keretből).

Elemzszám

A fenti mintavételi tervet 11200 elemű megvalósult mintára terveztük. A régiónkénti 1600-as elemzszám allokációja rétegek között a célsokaság nagyságával arányos módon történt.

10%-os kerethibát és 30%-os meghiúsulást feltételezve végül 17787 személyt választottunk ki.