

2. Импутация при немоделирана база

При тази група методи се използва намирането на донор от същото или предходно изследване независимо от вида на признака, който ще се импутира.

2.1. Донорна импутация.

Регресионната и пропорционалната импутация не могат да се използват във всички случаи. Когато характеристиките на единиците са на слабите скали, тези методи фактически са неприложими. В такива случаи може да се използва т. нар. донорна импутация. Този вид импутация, за разлика от разгледаните по-горе, се използва преди обработката на първичната информация. В случаи, когато единицата U има характеристики: X_1, X_2, X_3, X_4 и X_5 , където X_4 е липсваща, т. е. $(X_1, X_2, X_3, _, X_5)$, се търси подходяща единица донор. Единица, която има за 1, 2, 3 и 5 същите стойности $(X_1, X_2, X_3, Z_4, X_5)$, е търсеният донор и ние импутираме Z_4 на единицата U . При използването на този метод могат да възникнат проблеми: например при извадково проучване, в което анкетата има отговори и на интервалната, и евентуално на относителната скала освен на слабите скали. Когато и обемът на въпросите е голям, също е доста трудно да се намери донор със същите значения на признаците.

За единица № 7 за 1998 г., на която има нужда да се импутира стойност за средномесечен разход (табл. 1), донор не може да се намери, защото няма напълно съвпадение със значенията на признаците на нито една друга единица. Това донякъде се дължи на ограничения брой единици в примерното статистическо изследване.

2.2. "Hot desk"³ импутация

Този вид импутация има няколко разновидности. Той много прилича на донорната импутация, но всеки метод има своите специфични особености (Särndal, Swensson, Wretman, 1997, стр. 593). При него донорът винаги се определя от същата съвкупност по различни начини:

2.2.1. Случайна "hot desk" всеобхватна импутация. Донорът се избира от съвкупността случайно и донорните стойности се приписват на липсващите стойности. В конкретния случай трябва случайно да се избере донор от останалите единици в изследването за 1998 г. (табл. 1) и да се припише неговата стойност на единица № 7 за средномесечния разход.

³ Поради липса на терминология по темата на български език са оставени терминологичните понятия на английски език.