

си има своите особености, но понякога те се припокриват, защото при откриване на грешно въведени данни обикновено се импутира на тяхно място подходяща стойност. Процедурите, по които се избира стойност, и начинът, по който тя се импутира, представляват голям интерес от гледна точка на подобряване качеството на статистическите данни.

Импутацията на липсващи данни може да се раздели в зависимост от:

1. Виза на признака, чиито значения се импутират.

1.1. Признаци с количествени значения.

1.2. Признаци с качествени значения.

2. Начина, по който се импутират липсващите значения на признака.

2.1. Въз основа на донор (избира се по определен начин подходяща единица, която се явява донор, и с нейните характеристики се заместват (импутират) липсващите значения).

2.2. Въз основа на математическа или логическа процедура (например когато се намира модата при категорийни признаци) се намира характеристика, която се използва за импутация на липсващата стойност.

Обикновено импутацията не може да се разгледа като толкова опростена процедура. Трябва да се вземе предвид, че има припокриване между различните видове импутации. Затова описаното разделение на видовете импутации е условно и дава възможност да се класифицират до известна степен различни методи за импутация на липсващи данни.

За да се илюстрират по-добре методите за импутиране на липсващи данни, ще се опитам да приложа различни методи, като използвам резултати от примерно статистическо изследване. Ако при някои от методите се получават незадоволителни резултати, това не значи, че методът за импутация не може да се използва във всички случаи, а само че не е приложим в конкретния случай.

Нека вземем едно примерно статистическо изследване, при което има липсващи данни, на мястото на които ще се опитаме да импутираме по различните методи липсващите стойности.