

изследвания признак се анулират и чрез претегляне е възможно елиминиране на влиянието на неотговорилите. В този случай може да се говори за случайност при появата на липсващи стойности (MAR - missing at random) (Little, Rubin, 1987). Базовите тегла за единиците от извадка се дефинират като реципрочни на вероятността единицата да попадне в извадката. Има значение дали извадките са с равни вероятности за всички единици или не. Определянето на теглата за корекция на влиянието на неотговорилите се получава като базовите тегла се умножат по пропорцията на обема на групата и броя на отговорилите в групата. Основен проблем при претеглянето е определянето на групите.

Стандартният метод за третиране на проблеми от типа 3 е "**вмъкване**" (imputation) на стойности. Най-често използваните алгоритми са (Larsen, Madsen, 1999; Utsunomiya, Sonoda):

- **Метод на средната.** Използва се средното равнище като измерител на всички липсващи стойности. Прилага се както за цялата съвкупност, така и за отделни групи от нея.

- **Hot-Deck.** Липсващите стойности се заменят с такива, подбрани от специфична група респонденти с пълни отговори, които максимално се доближават по дадени характеристики до единицата, за която се налага "вмъкване".

- **Cold-Deck.** Липсващата стойност за дадена единица се заменя с постоянна, взета от външен източник. Оценки, получени от регресионни зависимости, и резултати от други изследвания също могат да бъдат използвани.

- **LVCF (Last, Value Carried Forward).** Последната регистрирана стойност се използва за попълване на липсващата следваща.

- **Multiple Imputation.** Този метод се разглежда от някои автори като подобрение на метода Hot-Deck. В основата си методът може да се представи като неколккратно повторение на метода Hot-Deck, като за основа за избор на стойности за вмъкване се използва подизвадка от единици от текущото изследване.

За защита от проблеми, описани в т. 4, се използва **корекция** (edit) на данните. Този процес може да бъде изпълняван "на ръка" или чрез електронна система. В своята същност той преминава през няколко етапа (Poirier, 2000, р. 1-11).

1. Определяне на правилата за корекция - идентифициране на съществуващите зависимости между включените в изследването променливи. Връзките могат да се представят като система от  $n$  линейни зависимости:

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m \leq b_1,$$

.....

$$a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nm}x_m \leq b_n,$$