

GIẢI ĐÁP THẮC MẮC

TẠI SAO PHẢI GIA QUYỀN SỐ LIỆU ĐIỀU TRA MẪU

BIẾN ĐỘNG DÂN SỐ VÀ KẾ HOẠCH HÓA GIA ĐÌNH NĂM 2000 VÀ 2001 ?

Phan Đắc Lộc
Vụ Dân số Lao động

Về mặt lý thuyết, khi thiết kế mẫu cho bất kỳ một cuộc điều tra mẫu thống kê nào đó, các chuyên gia mẫu rất phải lưu ý đến các yếu tố sau:

1/ Cấu trúc Dàn mẫu và đơn vị chọn mẫu cuối cùng;

2/ Xác định cỡ mẫu và tính đại diện của mẫu cho các chỉ tiêu cần nghiên cứu - thu thập từ mẫu;

3/ Phân bổ mẫu cho các tiểu vùng kinh tế - địa lý, các đơn vị hành chính, các khu vực thành thị, nông thôn, v.v...

4/ Lựa chọn phương pháp để chọn các đơn vị điều tra mẫu nhằm đảm bảo tính ngẫu nhiên, khách quan (chọn ngẫu nhiên lập và không lập với xác suất bằng nhau, chọn ngẫu nhiên hệ thống, chọn xác suất tỷ trọng, chọn mẫu phân tầng...).

5/ Xây dựng công thức tính: Độ lệch chuẩn, sai số mẫu, độ tin cậy;

6/ Gia quyền nội bộ số liệu điều tra mẫu theo **cơ cấu chuẩn** để tổng hợp số liệu qua từng cấp hoặc suy rộng số liệu điều tra mẫu theo số liệu tổng thể.

Trong thực tế khi sử dụng một dàn mẫu để chọn mẫu cho các cuộc điều tra về Biến động Dân số - KHHGD (BĐDS &

KHHGD), điều tra Lao động- việc làm hàng năm, v.v... chúng ta đều thấy rằng:

1/ Các đơn vị chọn mẫu có qui mô không bằng nhau, nhưng thực tế có thể nó lại được chọn ngẫu nhiên với xác suất bằng nhau, ví dụ:

- Đơn vị chọn điều tra mẫu là hộ gia đình: có hộ có qui mô lớn trên 15 khẩu, có hộ chỉ có 1-2 khẩu (Qui mô nhân khẩu trung bình khoảng 4,5 - 5 khẩu/1 hộ)

- Đơn vị chọn điều tra mẫu chòm là Địa bàn điều tra được phân định trong TĐTDS với qui mô chuẩn ở thành thị là 400 khẩu/ 1 ĐB; ở nông thôn là 500 khẩu/ 1 ĐB, Song trên thực tế có một số ĐB có qui mô quá lớn - trên 650 khẩu, và ngược lại có địa bàn có qui mô quá nhỏ dưới 200 khẩu.

- Đơn vị chọn điều tra mẫu là đơn vị hành chính xã, phường, hoặc là trường học, bệnh viện, cơ quan xí nghiệp... cũng có qui mô hoàn toàn khác nhau; Thực tế có xã trên 30000 dân, có xã chỉ có trên, dưới 200 dân.

2/ Tùy theo kinh phí cho phép, một mẫu được thiết kế điều tra có thể đủ lớn

để thu được số liệu đại diện cho cấp Toàn quốc, cấp tỉnh/thành phố, cấp huyện/quận, đại diện cho khu vực thành thị/nông thôn...

Do qui mô dân số của các tỉnh/ thành phố; các quận/ huyện cũng khác nhau nên việc phân bổ mẫu cho các địa phương thường không theo đúng tỷ trọng dân số. Nếu phân bổ theo tỷ trọng dân số thì một số tỉnh/thành phố lớn sẽ có một mẫu quá lớn (thừa mẫu); các tỉnh/thành phố nhỏ thì có một mẫu nhỏ không đủ đại diện (thiếu mẫu). Vì vậy các chuyên gia mẫu thường áp dụng việc phân tổ lại theo qui mô và xác định phạm vi mẫu tối thiểu cho từng tổ để đảm bảo tính đại diện.

Ví dụ:

PHÂN BỐ MẪU TRONG ĐIỀU TRA BDDS & KHHGD NĂM 2000 VÀ 2001:

Theo qui định trong phương án điều tra số 338/TCTK-DSLĐ ngày 27 tháng 4 năm 2000. Mẫu cho cuộc điều tra được thiết kế như sau:

"Trên cơ sở danh sách các địa bàn điều tra của cuộc Tổng điều tra dân số và nhà ở 1999. Vụ Dân số Lao động sẽ chọn ngẫu nhiên cho mỗi tỉnh/thành phố khoảng 24000 nhân khẩu (tương đương khoảng từ 47 đến 62 địa bàn điều tra, tùy theo qui mô địa bàn của từng tỉnh) để điều tra. Số địa bàn điều tra được phân bổ cho 2 khu vực thành thị và nông thôn theo tỷ trọng dân số thành thị và nông thôn trong từng tỉnh (theo số liệu TĐTDS và nhà ở 1999). Việc chọn địa bàn theo phương pháp ngẫu nhiên hệ thống riêng cho 2 khu vực".

Như vậy:

- Phạm vi điều tra mẫu trong nội bộ 1 tỉnh/thành phố, giữa 2 khu vực thành thị và nông thôn được xác định theo tỷ trọng, nhưng thực tế do qui mô địa bàn không giống nhau, do có sự thay đổi địa giới

hành chính, một số xã, năm 2000 đã được Nhà nước phê chuẩn thành phường. Do đó tỷ trọng thành thị/nông thôn mẫu và tỷ trọng dân số thành thị/nông thôn chuẩn của tỉnh/thành phố là không hoàn toàn giống nhau.

- Phạm vi điều tra mẫu của từng tỉnh/thành phố là xấp xỉ bằng nhau (tỷ trọng mẫu xấp xỉ nhau), nhưng thực tế tỷ trọng dân số chuẩn của một số tỉnh có sự chênh lệch khá nhiều, nhất là tỷ trọng dân số của các tỉnh mới tách (KonTum, Bắc Cạn...) lại là rất nhỏ.

Với việc phân bổ mẫu nêu trên, các đặc trưng thu được từ mẫu mang tính đại diện cho từng khu vực thành thị, nông thôn của từng tỉnh. Nếu tổng hợp trực tiếp các số liệu thu được từ mẫu trên mà không gia quyền theo tỷ trọng chuẩn thì số liệu của một số tỉnh/thành phố sẽ mang các đặc trưng nghiêng về đặc trưng của vùng thành thị (hoặc ngược lại, số liệu mẫu của một số tỉnh/ thành phố sẽ mang các đặc trưng nghiêng về nông thôn). Mặt khác, nếu tổng hợp trực tiếp các số liệu thu được từ mẫu trên mà không gia quyền theo tỷ trọng chuẩn thì số liệu của Toàn quốc sẽ mang các đặc trưng nghiêng về đặc trưng của các tỉnh có qui mô dân nhỏ.

Việc gia quyền số liệu điều tra mẫu theo cơ cấu chuẩn chính là việc hiệu chỉnh lại bức tranh số liệu thu được từ mẫu - là bức tranh thu nhỏ của tập dân số cho phù hợp với bức tranh chuẩn hiện tại.

CÁC HỆ SỐ GIA QUYỀN SỐ LIỆU ĐIỀU TRA MẪU BDDS & KHHGD NĂM 2000, 2001

Xuất phát từ việc thiết kế mẫu điều tra cho cuộc điều tra Biến động dân số và KHHGD năm 2000 và 2001, việc tổng hợp số liệu điều tra mẫu được gia quyền theo qui mô địa bàn TB, theo tỷ trọng thành thị/nông thôn chuẩn trong từng tỉnh, thành

phố và theo tỷ trọng dân số chuẩn của từng tỉnh, thành phố trong toàn quốc.

$$\boxed{W_{\text{tinh}} = W_1 * W_2}$$

Ở đây:

1/ Tổng hợp số liệu cho cấp tỉnh, thành phố:

Số liệu của các ĐBĐT mẫu được nhân với hệ số gia quyền **W tinh**:

- **W1**: Hệ số gia quyền theo quy mô nhân khẩu trung bình của ĐBĐT tỉnh riêng cho 2 khu vực thành thị và nông thôn.

$$W_1 = \frac{\text{Qui mô địa bàn chuẩn (hoặc qui mô nhân khẩu bình quân 1 ĐB)}}{\text{Qui mô từng địa bàn điều tra mẫu (số NK từng địa bàn mẫu)}}$$

- **W2**: Hệ số gia quyền theo tỷ trọng TT/NT chuẩn của từng tỉnh, thành phố (Lấy theo báo cáo chính thức DS TB năm 2000 và 2001).

$$W_2 = \frac{\text{Tỷ trọng dân số Thành thị hoặc Nông thôn chuẩn (số liệu toàn bộ)}}{\text{Tỷ trọng dân số Thành thị hoặc Nông thôn điều tra mẫu}}$$

2/ Tổng hợp số liệu cho 8 vùng và Toàn quốc:

Số liệu của các ĐBĐT mẫu được nhân với hệ số gia quyền **WTQuoc**:

$$\boxed{WT_{\text{Quoc}} = W_1 * W_2 * W_3}$$

Ở đây: **W3**: Hệ số gia quyền theo tỷ trọng dân số chuẩn của từng tỉnh, thành phố trong toàn quốc, (lấy theo báo cáo chính thức DS TB 2000 và năm 2001).

$$W_3 = \frac{\text{Tỷ trọng dân số toàn bộ của từng tỉnh/thành phố (so toàn quốc)}}{\text{Tỷ trọng dân số điều tra mẫu của từng tỉnh/thành phố (so Tổng số DS mẫu TQ)}}$$