

II. THIẾT KẾ VÀ ƯỚC LƯỢNG MẪU

1. Dàn chọn mẫu

Mẫu của cuộc Điều tra lao động và việc làm năm 2010 là mẫu hệ thống phân tầng hai cấp, đại diện cho cấp tỉnh. Mỗi tỉnh, thành phố trực thuộc Trung ương tạo thành một tầng chính với hai tầng thứ cấp là khu vực thành thị và khu vực nông thôn. Dàn chọn mẫu là các địa bàn điều tra mẫu 15% của Tổng điều tra dân số và nhà ở 2009.

2. Xác định cỡ mẫu và phân bổ mẫu

Để đảm bảo thu được các ước lượng mẫu đại diện cho cấp tỉnh, thành phố, mẫu được phân bổ theo phương pháp tỷ lệ nghịch với quy mô dân số, mỗi tỉnh đã được chọn khoảng 60 địa bàn với quy mô 32 hộ/địa bàn. Hà Nội và thành phố Hồ Chí Minh là hai thành phố đông dân, phức tạp nhất cả nước đã được chọn 74 địa bàn (Phụ lục 2). Cỡ mẫu: bình quân 2.000 hộ/tỉnh mỗi kỳ. Mỗi kỳ của cuộc điều tra lao động và việc làm năm 2010 đã được tiến hành ở 3890 địa bàn mẫu trên phạm vi cả nước. Tổng cục Thống kê đã trực tiếp chọn các địa bàn điều tra và thông báo cho các Cục Thống kê.

Mẫu của cuộc điều tra là mẫu phân tầng, mỗi tỉnh/thành phố tạo thành một tầng chính (63 tầng chính), sau đó được tách riêng theo hai tầng thứ cấp là “thành thị” và “nông thôn”. Việc phân bổ mẫu cho mỗi tầng được thực hiện theo phương pháp chọn hệ thống. Các địa bàn mẫu đã chọn trong bước chuẩn bị, Vụ Thống kê Dân số và Lao động tiến hành chọn hệ thống cấp 1 là địa bàn điều tra, cấp 2: chọn ngẫu nhiên 32 hộ từ danh sách hộ của địa bàn điều tra.

3. Ước lượng mẫu

3.1 Ước lượng mẫu cho 1 kỳ điều tra

Quyền số chung có thể được tính toán dựa vào xác suất/quyền số sau:

- 1) Quyền số thiết kế (quyền số cơ bản): dựa vào xác suất;
- 2) Hệ số hiệu chỉnh quyền số do thay đổi số hộ hoặc thay đổi tổng số địa bàn do mất đi mà không chọn thay thế;

3) Hệ số hiệu chỉnh quyền số theo cơ cấu tổng thể nghiên cứu (gia quyền).

Ký hiệu:

- W_{1hji} - Quyền số thiết kế (quyền số cơ bản) của địa bàn j , tầng h ;
- W_{2hji} - Hệ số hiệu chỉnh quyền số do số hộ (dân số) thay đổi;
- W_{3hji} - Hệ số hiệu chỉnh quyền số theo quy mô địa bàn trung bình của tầng h ;
- W_{4hji} - Hệ số hiệu chỉnh quyền số do số địa bàn điều tra thay đổi;
- W_{5hji} - Hệ số hiệu chỉnh quyền số theo cơ cấu tổng thể nghiên cứu;
- W_{hji} - Quyền số mẫu đối với hộ hoặc dân số nam/nữ của địa bàn j tầng h .

Xác định quyền số cơ bản

Theo thiết kế, cuộc điều tra Lao động và Việc làm được tiến hành trên phạm vi mẫu xấp xỉ như nhau ở tất cả 63 tỉnh, thành phố trong cả nước. Mỗi tỉnh, thành phố điều tra khoảng 24.000 hộ gia đình.

Giả sử a_h là số địa bàn điều tra được chọn trong tầng h và N_h là tổng số địa bàn của tầng h . Do mẫu được chọn độc lập ở từng tầng theo phương pháp ngẫu nhiên hệ thống, nên xác suất chọn cơ bản được tính theo công thức sau: $P_{1hji} = \frac{a_h}{N_h}$ và quyền số cơ bản (quyền số thiết kế) của địa bàn j thuộc tầng h là nghịch đảo của xác suất chọn, được tính như sau:

$$W_{1hji} = \frac{1}{P_{1hji}} = \frac{N_h}{a_h} \approx \frac{M_h}{\sum m_{hj}}$$

Trong đó, M_h là tổng số hộ (dân số) của tầng h và $\sum m_{hj}$ là tổng số hộ (dân số) của các địa bàn đã chọn điều tra của tầng h .

Xác định hệ số hiệu chỉnh quyền số do thay đổi số hộ (dân số) và số địa bàn

a) Hiệu chỉnh quyền số do thay đổi số hộ (dân số):

Giả sử m_{hj} là tổng số hộ (dân số) khi lập bảng kê của địa bàn j của tầng h và m_{hj}^* là tổng số hộ (dân số) khi điều tra của địa bàn j của tầng h . Do các địa bàn trong từng tầng được chọn với xác suất như nhau và được chọn từ dân

mẫu của Tổng điều tra dân số và nhà ở năm 2009 và hệ số hiệu chỉnh do thay đổi số hộ (dân số) được tính theo công thức sau:

$$W_{2hji} = \frac{1}{P_{2hji}} = \frac{m_{hj}}{m_{hj}^*}$$

Các địa bàn của Tổng điều tra dân số và nhà ở năm 2009 được phân chia với quy mô không đều nhau, nên cần phải xác định hệ số điều chỉnh quy mô hộ/dân số của các địa bàn Tổng điều tra dân số và nhà ở năm 2009 về quy mô hộ/dân số trung bình của tầng đó. Giả sử $\overline{m}_{hj}$ là tổng số hộ (dân số) trung bình của địa bàn thuộc tầng h và hệ số hiệu chỉnh do thay đổi số hộ (dân số) được tính theo công thức sau:

$$W_{3hji} = \frac{1}{P_{3hji}} = \frac{\overline{m}_{hj}}{m_{hj}}$$

b) Hiệu chỉnh quyền số do thay đổi số địa bàn:

Theo quy định: nếu địa bàn nào đã được chọn mà trong quá trình hiệu chỉnh sơ đồ - bảng kê phát hiện đã bị giải toả hoặc mất đi thì được phép thay thế bằng 1 địa bàn liền kề, không thay đổi tổng số địa bàn đã được chọn. Nên:

$$W_{4hji} = \frac{1}{P_{4hji}} = 1$$

Xác định hệ số hiệu chỉnh quyền số theo cơ cấu tổng thể (gia quyền theo tỷ trọng dân số nghiên cứu)

Dân số hàng năm được ước lượng theo thành thị/nông thôn và giới tính cho 63 tỉnh/thành phố, nên có thể gia quyền theo tỷ trọng dân số thành thị/nông thôn và dân số nam/nữ. Giả sử m_{hji}^* là tổng số hộ (dân số nam/nữ) khi điều tra của địa bàn j tầng h ; $m_{hji}^{*'}$ là tổng số hộ (dân số nam/nữ) hiệu chỉnh theo tỷ trọng thành thị/nông thôn và nam nữ của địa bàn j tầng h và tính theo công thức:

$$m_{hji}^{*'} = m_{hji}^* \times \frac{M_{hi}^*}{M_h^*}$$

Trong đó:

m_{hj}^* số hộ (dân số nam/nữ) thu được từ điều tra mẫu của địa bàn j tầng h ;

M_{hi}^* số hộ (dân số nam/nữ) chia theo thành thị/nông thôn ước đến 1/4/2010 (hoặc 1/10/2010) của tầng h ; ($i = 1$ – thành thị ; $i = 2$ – nông thôn)

M_h^* số hộ (dân số) ước đến 1/4/2010 (hoặc 1/10/2010) của tầng h .

Hệ số hiệu chỉnh theo cơ cấu tổng thể của dân số (số hộ) ước lượng đến 1/4/2010 (hoặc 1/10/2010) được xác định như sau :

$$W_{shji} = \frac{1}{P_{shji}} = \frac{m_{hji}^*}{m_{hji}^*} \times \frac{M_h^*}{M_h} = \frac{m_{hj}^*}{m_{hji}^*} \times \frac{M_{hi}^*}{M_h^*} \times \frac{M_h^*}{M_h} = \frac{m_{hj}^*}{m_{hji}^*} \times \frac{M_{hi}^*}{M_h}$$

Vì phân bố mẫu là không tỷ lệ thuận đối với các tổng thể nghiên cứu, nên các quyền số mẫu sẽ được tính cho tất cả các phân tích sử dụng số liệu của điều tra nhằm đảm bảo tính đại diện thực tế của mẫu. Quyền số mẫu đối với mỗi hộ (hoặc dân số loại i) của địa bàn j thuộc tầng h là nghịch đảo của xác suất chọn:

$$W_{hji} = 1/P_{hji} = W_{1hji} \times W_{2hji} \times W_{3hji} \times W_{4hji} \times W_{5hji}$$

$$W_{hji} = 1/P_{hji} = \frac{M_h}{\sum m_{hj}} \times \frac{m_{hj}}{m_{hj}^*} \times \frac{\bar{m}_{hj}}{m_{hj}} \times \frac{m_{hj}^*}{m_{hji}^*} \times \frac{M_{hi}^*}{M_h} = \frac{\bar{m}_{hj}}{\sum m_{hj}} \times \frac{M_{hi}^*}{m_{hji}^*}$$

Vụ Thống kê Dân số - Lao động đã phối hợp chặt chẽ với Trung tâm Tin học Thống kê Khu vực I lập trình, tính toán cụ thể và kiểm tra chính xác các quyền số trên cho tất cả 3890 địa bàn của cuộc điều tra.

3.2 Ước lượng mẫu trung bình cho các kỳ điều tra

Điều tra Lao động và Việc làm năm 2010 được tiến hành vào 2 thời điểm (kỳ 1: 1/4/2010 và kỳ 2: 1/10/2010). Để gia quyền số liệu cho năm 2010, phải tính quyền số chung cho cả 2 kỳ điều tra, theo công thức sau:

$$W_{hji}^* = \frac{W_{hji}^1 + W_{hji}^2}{2}$$

Trong đó:

- W_{hji}^* - Quyền số cả năm đối với hộ hoặc dân số nam/nữ của địa bàn j tầng h .
- W_{hji}^1 - Quyền số kỳ 1 đối với hộ hoặc dân số nam/nữ của địa bàn j tầng h .
- W_{hji}^2 - Quyền số kỳ 2 đối với hộ hoặc dân số nam/nữ của địa bàn j tầng h .