

Một số ý kiến về phương pháp phức tra trong Tổng điều tra dân số và nhà ở

Nguyễn Văn Phái

Trong các cuộc Tổng điều tra dân số và nhà ở các năm 1989 và 1999, Ban Chỉ đạo TĐTDS và nhà ở Trung ương đã sử dụng phương pháp hệ thống đôi hay song trùng (Dual Record) khi phức tra để đánh giá mức độ đầy đủ của công tác điều tra ghi phiếu. Từ việc so sánh thông tin của phiếu điều tra và phiếu phức tra nhằm có được ước lượng mức độ sai sót. Sau khi tính các tỷ lệ sai sót, kết quả phức tra cho thấy khoảng 1,3 phần trăm dân số đã bị bỏ sót trong Tổng điều tra. Tỷ lệ đó là thấp theo tiêu chuẩn chung của thế giới và cho phép tin tưởng vào giá trị Tổng điều tra ¹.

Tuy nhiên, cần phải nhận thấy rằng, phương pháp hệ thống đôi không phù hợp cho việc đánh giá mức độ chính xác của số liệu Tổng điều tra dân số và nhà ở.

Về lý thuyết, phương pháp hệ thống đôi được xây dựng để đánh giá mức độ bỏ sót trong đăng ký các sự kiện (biến cố) nhân khẩu học đã xảy ra, chủ yếu là các sự kiện hộ tịch (sinh, chết, kết hôn). Việc đăng ký các sự kiện hộ tịch chỉ có thể xảy ra khả năng bỏ sót (chưa/không đăng ký) chứ không thể xảy ra khả năng đăng ký trùng (đăng ký 2 lần), do khi tiến hành đăng ký các sự kiện hộ tịch, bắt buộc phải có tài liệu khẳng định sự kiện đó xảy ra (giấy chứng sinh, chứng tử...) mà những tài liệu đó không được cấp 2 lần cho một sự kiện. Bởi vậy, sau khi công tác điều tra các sự kiện hộ tịch xảy ra trong năm được thực

hiện một cách độc lập, người ta tiến hành so sánh các sự kiện đã đăng ký với những sự kiện hộ tịch đã điều tra được. Khi tiến hành so sánh, có thể phân loại số các sự kiện theo các dạng sau:

A1 - số sự kiện hộ tịch có cả trong đăng ký lẫn trong điều tra;

A2 - số sự kiện hộ tịch chỉ thu được trong đăng ký (bỏ sót trong điều tra);

A3 - số sự kiện hộ tịch chỉ thu được trong điều tra (bỏ sót trong đăng ký).

Theo lý thuyết xác suất, số sự kiện hộ tịch có thể bỏ sót trong cả trong đăng ký lẫn trong điều tra sẽ là:

$$\frac{A2 \times A3}{A1}$$

Như vậy, số sự kiện hộ tịch thực tế xảy ra trong năm trên địa bàn sẽ là:

$$A = A1 + A2 + A3 + \frac{A2 \times A3}{A1}$$

Từ đó có thể xác định được tỷ lệ bỏ sót (μ) trong đăng ký sự kiện hộ tịch đó sẽ là:

$$\mu = \frac{A2 + A3}{A} \times 100$$

¹ Ban chỉ đạo Tổng điều tra dân số và nhà ở Trung ương. Tổng điều tra dân số và nhà ở 1999. KẾT QUẢ ĐIỀU TRA MẪU. Nhà xuất bản Thế giới, Hà Nội, 2000. tr.24

Danh sách 1. Các hộ và nhân khẩu trong hộ đã được đăng ký trong điều tra		
Họ tên chủ hộ	Họ tên các nhân khẩu được điều tra	Ghi chú
Nguyễn Văn X	1. Nguyễn Văn X 2. Trần Thị B 3. Nguyễn Văn C	Chủ hộ Vợ Con trai (Sỹ quan BCH QS tỉnh)
Lê Thị Y	1. Lê Thị Y 2. Hồ Văn D	Chủ hộ Chồng
Trần Văn Z	1. Trần Văn Z 2. Trịnh Thị E 3. Trần Văn G 4. Trần Thị H	Chủ hộ Vợ Con trai Con gái
Vũ Văn I	1. Vũ Văn I 2. Nguyễn Thị K 3. Vũ Thị L 4. Vũ Văn M	Chủ hộ Vợ Con gái Con trai

Danh sách 2. Các hộ và nhân khẩu trong hộ đã được đăng ký trong phức tra		
Họ tên chủ hộ	Họ tên các nhân khẩu được điều tra	Ghi chú
Nguyễn Văn X	1. Nguyễn Văn X 2. Trần Thị B 3. Nguyễn Thị O	Chủ hộ Vợ Con gái đi xuất khẩu lao động còn thời hạn
Lê Thị Y	1. Lê Thị Y 2. Hồ Văn D 3. Hồ Văn P	Chủ hộ Chồng Con, học sinh phổ thông đi trợ học
Trần Văn Z	1. Trần Văn Z 2. Trịnh Thị E 3. Trần Văn G 4. Trần Thị H	Chủ hộ Vợ Con trai Con gái
Vũ Văn I	1. Vũ Văn I 2. Nguyễn Thị K 3. Vũ Thị L 4. Vũ Văn M	Chủ hộ Vợ Con gái Con trai

Trong khi đó, trong tổng điều tra, một nhân khẩu nào đó có thể bị bỏ sót trong đăng ký (không có nơi nào đăng ký) nhưng cũng có thể có nhân khẩu bị đăng ký trùng (được đăng ký ở 2 nơi). Nói một cách khác, trong tổng điều tra dân số khả năng ghi trùng và bỏ sót nhân khẩu đều có thể xảy ra. Bởi vậy, nếu sử dụng phương pháp hệ thống đôi, sẽ không thể xác định được tỷ lệ ghi trùng, thậm chí, tỷ lệ bỏ sót có thể bị tăng lên so với thực tế do những nhân khẩu ghi trùng sẽ được tính thành bỏ sót. Ví dụ dưới đây sẽ chứng minh khi sử dụng phương pháp hệ thống đôi, tỷ lệ bỏ sót thu được qua phức tra bị cao hơn so với thực tế:

Giả sử Địa bàn X có 4 hộ (giả định), trong bước điều tra, điều tra viên đã xác định được số nhân khẩu trong từng hộ như sau (**Danh sách 1**):

Khi tiến hành phức tra, cán bộ phức tra, qua phỏng vấn trực tiếp theo qui trình như đã sử dụng trong tổng điều tra đã thu được số nhân khẩu trong từng hộ như sau (**Danh sách 2**):

Theo phương pháp hệ thống đôi, khi so sánh hai bản danh sách trên, chúng ta có số các trường hợp A1, A2, A3 như sau:

A1 (có trong cả hai danh sách) - 12 người: Nguyễn Văn X, Trần Thị B (hộ ông X); Lê Thị Y, Hồ Văn D (hộ bà Y); Trần Văn Z, Trịnh Thị E, Trần Văn G, Trần Thị H (hộ ông Z); Vũ Văn I, Nguyễn Thị K, Vũ Thị L, Vũ Văn M (hộ ông I);

A2 (chỉ có trong danh sách 2 – phức tra) – 2 người: Nguyễn Thị O (hộ ông X) và Hồ Văn P (hộ bà Y);

A3 (chỉ có trong danh sách 1 - điều tra) – 1 người: Nguyễn Văn C (hộ ông X).

A (tổng số nhân khẩu sẽ là):

$$1 \times 2 \\ = 12 + 2 + 1 + \frac{1 \times 2}{12} = 15,17$$

Theo phương pháp song trùng, tỷ lệ bỏ sót trong điều tra là:

$$\mu = \frac{15,17 - 12}{15,17} \times 100 = 20,9\%$$

Tuy nhiên, trong thực tế, theo qui định đối tượng điều tra ghi phiếu tại địa bàn này chỉ bỏ sót 2 nhân khẩu, đó là Nguyễn Thị O (hộ ông X) và Hồ Văn P (hộ bà Y) là các đối tượng điều tra ghi phiếu tại hộ nhưng chưa được đăng ký. Anh Nguyễn Văn C mặc dù sinh sống tại hộ nhưng theo qui định sẽ do Ban Chỉ đạo TĐTDS Bộ Quốc phòng điều tra nên sẽ là trường hợp đăng ký trùng. Như vậy, nếu điều tra đúng, địa bàn này sẽ chỉ có 14 nhân khẩu. Theo đó, tỷ lệ bỏ sót (μ) và ghi trùng (α) thực tế của cuộc điều tra sẽ là:

$$\mu = 2/14 * 100 = 14,2\%$$

$$\alpha = 1/14 * 100 = 7,2\%$$

Tương ứng theo đó, tỷ lệ sai sót thực tế của điều tra sẽ là:

$$\alpha - \mu = 7,1\% - 14,2\% = - 7,1\%$$

Rõ ràng, con số sai số thực tế là 7,1% thấp hơn nhiều so với con số 20,9% tính được theo

phương pháp hệ thống đôi.

Kết luận

Phương pháp hệ thống đôi được xây dựng rất khoa học dựa trên cơ sở lý thuyết xác suất và thống kê toán. Tuy nhiên, phương pháp này chỉ có thể sử dụng việc đánh giá mức độ đầy đủ (bỏ sót) trong công tác đăng ký các sự kiện hộ tịch mà không thể sử dụng cho phức tra nhằm đánh giá mức độ đầy đủ trong đăng ký nhân khẩu trong TĐTDS. Thậm chí, phương pháp hệ thống đôi cũng không thể sử dụng để đánh giá mức độ trùng sót về các trường hợp sinh chết trong các cuộc điều tra biến động dân số và kế hoạch hoá gia đình hàng năm vì trong cuộc điều tra chọn mẫu, các trường hợp sinh chết được thu thập dựa trên cơ sở phỏng vấn nên vẫn có khả năng khai trùng các sự kiện sinh chết.

Trong TĐTDS, khi phức tra để đánh giá mức độ chính xác về số lượng nhân khẩu trong điều tra có thể sử dụng phương pháp điều tra lặp. Cán bộ phức tra sau khi được huấn luyện nghiệp vụ tiến hành phỏng vấn lại tất cả các hộ thuộc địa bàn phức tra. Trên cơ sở so sánh 2 danh sách (điều tra và phức tra) xem những nhân khẩu nào đúng là đối tượng điều tra ghi phiếu tại địa bàn, những nhân khẩu nào thuộc loại ghi trùng (không phải là đối tượng điều tra), những nhân khẩu nào bị bỏ sót để từ đó xác định cụ thể các tỷ lệ trùng (α), sót (μ) và sai số ($\alpha - \mu$) của cuộc tổng điều tra. ■

Tính chỉ số khối lượng sản phẩm công nghiệp...

(Tiếp theo trang 17)

triệu đồng và năm 2008 là 8.779 nghìn đồng.

Với các số liệu đó ta tính được các hệ số thay đổi mặt hàng sản xuất như sau:

- Năm 2007: $K_0 = \frac{8960}{8960 - 1250} = 1,1621$

- Năm 2008: $K_1 = \frac{8779}{8779 - 2200} = 1,3344$

Từ đó ta có: $I_K = \frac{1,3344}{1,1621} = 1,1482$

Với kết quả tính toán $I_K = 1,1482$ và các chỉ số khối lượng sản phẩm so sánh được theo công thức 1b ở trên: $I'_q = 0,7908$; tiếp tục tính được chỉ số khối lượng sản phẩm:

$I_q = 0,7908 \times 1,1482 = 0,9080$ hoặc 90,80%

Số liệu trên cho thấy: Nếu áp dụng phương pháp tính chỉ số khối lượng sản phẩm dạng bình quân, nhưng có tính đến yếu tố mở rộng hay thu hẹp mặt hàng thông qua chỉ số I_K như đã trình bày thì chỉ số khối lượng sản phẩm có kết quả là 0,9080, cao hơn chỉ số chỉ tính trên các mặt hàng so sánh được là 0,1172 (0,9080 - 0,7908).

Trên đây là ví dụ để làm rõ nội dung và phương pháp tính cũng như cách vận dụng của phương pháp tính chỉ số khối lượng sản phẩm công nghiệp. Trong thực tế, việc tính giá trị sản phẩm không so sánh được theo giá thực tế một cách thường xuyên ở tất cả các cấp độ: doanh nghiệp, ngành công nghiệp riêng biệt và toàn ngành công nghiệp không đơn giản. Cần phải tiếp tục nghiên cứu để xác định xu thế tăng lên của yếu tố mở rộng mặt hàng làm căn cứ tính toán chỉ số khối lượng sản phẩm được thuận lợi. Vấn đề này chúng tôi sẽ giới thiệu ở các số báo sau. ■