

Đánh giá chất lượng số liệu điều tra thực trạng việc làm và thất nghiệp năm 2007

TS. Phạm Đăng Quyết ^(*)

Từ “chất lượng” có nhiều nghĩa khác nhau phụ thuộc vào nội dung mà nó được sử dụng. Ở nghĩa rộng, chất lượng số liệu có thể được định nghĩa trên giác độ đầu ra thống kê tốt như thế nào khi được phổ biến – có đáp ứng được nhu cầu và kỳ

vọng của người dùng tin đối với thông tin thống kê không. Nghĩa hẹp, chất lượng được đồng nhất với chính xác thống kê. Để đo lường chất lượng số liệu, Thống kê Châu Âu đưa ra 6 tiêu thức đánh giá tại bảng sau:

Định nghĩa	Thành tố chủ yếu
1. Thích hợp	
Mức độ mà sản phẩm thống kê đáp ứng nhu cầu của người sử dụng trên cả hai phương diện phạm vi và nội dung	Mọi sự đánh giá tính thích hợp cần xét tới: - Ai là người sử dụng số liệu thống kê; - Họ cần cái gì; và - Đầu ra đáp ứng như thế nào đối với các nhu cầu đó?
2. Chính xác	
Mức độ gần giữa kết quả ước tính và giá trị thực (không biết)	Tính chính xác có thể được chia ra thành sai số chọn mẫu và sai số phi chọn mẫu, sai số phi chọn mẫu bao gồm: - Sai số phạm vi; - Sai số không trả lời; - Sai số đo lường; - Sai số xử lý; - Sai số giả định mô hình.
3. Kịp thời và đúng thời gian	
Tính kịp thời có thể quy tới khoảng thời gian giữa công bố và chu kỳ thu thập số liệu. Đúng thời gian có thể quy cho thời gian khoảng giữa ngày công bố thực tế với ngày công bố theo kế hoạch.	Sự đánh giá tính kịp thời và đúng thời gian có tính đến những điều sau: - Thời gian sản xuất số liệu; - Tần suất phát hành; và - Đúng thời gian phát hành.

^[*] Bộ Lao động – Thương binh và Xã hội

Định nghĩa	Thành tố chủ yếu
4. Tiếp cận và rõ ràng	
<p>Sự tiếp cận là việc những người sử dụng có thể tiếp cận tới số liệu dễ dàng. Nó cũng liên quan tới khổ sách mà số liệu có và tính có thể dùng được của các thông tin hỗ trợ. Tính rõ ràng có thể quy tới chất lượng và tính đầy đủ của siêu dữ liệu, minh hoạ và sự chỉ dẫn đi kèm.</p>	<ul style="list-style-type: none"> - Nhu cầu phân tích; - Sự trợ giúp xác định thông tin; - Tính rõ ràng; và - Phổ biến số liệu.
5. So sánh	
<p>Mức độ mà số liệu có thể so sánh theo thời gian và phạm vi lĩnh vực.</p>	<p>Tính so sánh có thể được xem trên giác độ so sánh về:</p> <ul style="list-style-type: none"> - Thời gian; - Không gian (ví dụ: tỉnh, vùng, cả nước); và - Phạm vi/ lĩnh vực hay bộ phận của tổng thể (ví dụ: ngành, loại hộ).
6. Chặt chẽ	
<p>Mức độ mà số liệu được thu thập từ các nguồn hay các phương pháp khác nhau nhưng có thể quy đến hiện tượng giống nhau, tương tự.</p>	<p>Tính chặt chẽ được xem xét trên giác độ chặt chẽ giữa:</p> <ul style="list-style-type: none"> - Số liệu được sản xuất ở các tần suất khác nhau; - Số liệu thống kê khác trong cùng phạm vi/lĩnh vực kinh tế - xã hội; và - Các nguồn số liệu và đầu ra.

Đảm bảo chất lượng được thực hiện theo các giai đoạn của quá trình sản xuất số liệu thống kê: từ thiết kế, thu thập dữ liệu, xử lý số liệu, quyền số và ước tính, phân tích số liệu, kiểm tra tính rõ ràng thống kê, đến phổ biến số liệu. Thiết lập và duy trì chất lượng số liệu là cần thiết đối với sự thành công của hệ thống thống kê.

Bộ Lao động – Thương binh và Xã hội tiến hành cuộc điều tra thực trạng việc làm và thất nghiệp nhằm thu thập thông tin phản ánh thực trạng việc làm và thất nghiệp trên thị trường lao động của các tỉnh, thành phố, các vùng kinh tế trọng điểm, các vùng lãnh thổ và cả nước, đặc biệt là tình hình thất nghiệp khu vực thành thị và thiếu việc làm khu vực nông thôn để phục vụ việc triển khai thực hiện các chính sách và giải pháp về phát triển thị trường lao động của Chính phủ và các cấp, các ngành.

Phạm vi điều tra tại 64 tỉnh/thành phố trong toàn quốc. Quy mô mẫu điều tra là 100.680 hộ gia đình với số lượng địa bàn mẫu là 3.356 địa bàn. Đơn vị điều tra là hộ gia đình thuộc các địa bàn điều tra đã được chọn. Số địa bàn điều tra và hộ điều tra được phân bổ, chọn mẫu theo khu vực thành thị – nông thôn của các tỉnh, thành phố. Đối tượng điều tra là nhân khẩu thực tế thường trú từ đủ 15 tuổi trở lên.

Để ước lượng suy rộng các chỉ tiêu cơ bản thu được từ cuộc Điều tra thực trạng việc làm và thất nghiệp, phương pháp suy rộng dùng hệ số nhân (quyền số) là tỷ lệ mẫu nghịch đảo dựa vào số liệu báo cáo dân số trung bình hàng năm (dân số có đến 1/7) do Tổng cục Thống kê công bố được sử dụng theo quy trình sau:

- Tính các quyền số suy rộng rồi tiến hành suy rộng cho từng địa bàn;

- Tính quyền số suy rộng trung bình cho tất cả các địa bàn thuộc mỗi tầng thứ cấp (thành thị, nông thôn), sau đó suy rộng riêng cho thành thị, nông thôn của từng tỉnh;

- Dựa trên kết quả suy rộng cho thành thị và nông thôn, tiến hành suy rộng chung cho toàn tỉnh/ thành phố.

Theo phương pháp này, có 256 hệ số nhân (64 tỉnh x 4) được tính như sau:

$f^1 = S^1/m^1$	ở đây: S^1 -Tổng số nam thuộc khu vực thành thị của tỉnh có đến 1/7 m^1 - Tổng số nam thuộc mẫu thành thị của tỉnh
$f^2 = S^2/m^2$	ở đây: S^2 -Tổng số nam thuộc khu vực nông thôn của tỉnh có đến 1/7 m^2 - Tổng số nam thuộc mẫu nông thôn của tỉnh
$f^3 = S^3/m^3$	ở đây: S^3 -Tổng số nữ thuộc khu vực nông thôn của tỉnh có đến 1/7 m^3 - Tổng số nữ thuộc mẫu nông thôn của tỉnh
$f^4 = S^4/m^4$	ở đây: S^4 -Tổng số nữ thuộc khu vực thành thị của tỉnh có đến 1/7 m^4 - Tổng số nữ thuộc mẫu thành thị của tỉnh.

Trong phương án điều tra có quy định kiểm tra chất lượng số liệu điều tra nhằm giảm thiểu sai số phi chọn mẫu bao gồm những bước sau:

1. Việc điều tra ghi phiếu phải thực hiện theo nguyên tắc phỏng vấn trực tiếp đủ số lượng các hộ thuộc địa bàn điều tra, thực hiện đúng những quy định về nghiệp vụ điều tra; không bỏ sót đối tượng điều tra, không ghi sai hoặc bỏ sót những nội dung đã quy định trong phiếu điều tra.

2. Tổ trưởng điều tra phải kiểm tra toàn bộ những thông tin mà điều tra viên đã ghi chép trong phiếu hộ, nếu phát hiện những sai sót phải yêu cầu điều tra viên trở lại hộ để xác minh và sửa phiếu điều tra đảm bảo độ tin cậy của các thông tin ghi trên phiếu. Tuyệt đối không được tự ý sửa phiếu.

3. Giám sát viên có nhiệm vụ giám sát toàn bộ quy trình ghi phiếu tại địa bàn được phụ trách của điều tra viên và tổ trưởng điều tra. Khi phát hiện sai sót phải thông báo cho điều tra viên, tổ trưởng điều tra để kịp thời sửa chữa.

4. Ban chỉ đạo điều tra các tỉnh/thành phố có trách nhiệm tạo điều kiện thuận lợi và phối hợp chặt chẽ với Ban chỉ đạo điều tra trung ương, lực lượng cán bộ kiểm tra trung ương tiến hành kiểm tra tình hình triển khai thực hiện điều tra trên địa bàn. Nếu chất lượng điều tra không đạt yêu cầu một cách có hệ thống, Ban chỉ đạo điều tra Trung ương sẽ quyết định tổ chức điều tra lại và Ban chỉ đạo điều tra các tỉnh/thành phố phải hoàn toàn chịu trách nhiệm giải quyết phần kinh phí phát sinh có liên quan.

Song trên thực tế, công tác làm sạch, kiểm tra phiếu trước khi nhập tin, kiểm tra dữ liệu gốc và hiệu đính với sự hỗ trợ của máy tính đòi hỏi rất nhiều thời gian. Điều tra được thực hiện tại thời điểm 1/7/2007, nghiệm thu phiếu điều tra vào tháng 8/2007 và đến tháng 7/2008 mới phát hành được kết quả điều tra.

Để có được đánh giá về chất lượng số liệu điều tra, chúng ta sử dụng chương trình STATA tính toán các sai số điều tra cho chỉ tiêu tỷ lệ thất nghiệp của cả nước, thành thị và nông thôn, tính toán hệ số biến sai so sánh độ chính xác của các ước lượng này.

```
Survey: Ratio estimation
Number of strata = 1           Number of obs = 323564
Number of PSUs = 323564      Population size = 6.7e+07
                               Design df = 323563
    _ratio_1: thatnghiep/lucluonglaodong
    _subpop_1: ttnt = thành thị
    _subpop_2: ttnt = nông thôn
```

Over	Ratio	Linearized Std. Err.	[95% Conf. Interval]	
_ratio_1	.0241752	.0004209	.0233501	.0250002
_subpop_1	.0491259	.0009899	.0471857	.051066
_subpop_2	.0156844	.0004489	.0148045	.0165642

```
. svyset _n [pweight=quyenso], vce(linearized)
```

Kết quả tính toán cho thấy: tỷ lệ thất nghiệp của cả nước năm 2007 ước tính là 2,41% với sai số chuẩn là 0,04%; tỷ lệ thất nghiệp khu vực thành thị là 4,91% với sai số chuẩn là 0,09%, tỷ lệ thất nghiệp khu vực nông thôn là 1,56% với sai số chuẩn là 0,04%.

Độ chính xác hay độ lớn của sai số từ điều tra chọn mẫu được thể hiện ở phương sai và độ lệch chuẩn của nó. Sai số chuẩn là căn bậc hai của phương sai mẫu, nó đại diện cho biến thiên ngẫu nhiên của một ước lượng do thiết kế mẫu riêng của nó. Ta thấy tỷ lệ thất nghiệp khu vực thành thị có sai số (0,09%) lớn hơn nhiều so với sai số tỷ lệ thất nghiệp của cả nước và của khu vực nông thôn (0,04%).

Song để so sánh độ chính xác của một ước lượng so với một ước lượng khác chúng ta sử dụng hệ số biến sai. Hệ số biến sai (CV) hay còn gọi là sai số chuẩn tương đối (RSE) được thể hiện dưới dạng phần trăm của giá trị sai số chuẩn so với giá trị ước lượng:

$$CV(\bar{y}) = \frac{SE(\bar{y})}{\bar{Y}} \times 100$$

Theo công thức trên ta có giá trị hệ số biến sai tỷ lệ thất nghiệp của cả nước, thành thị và nông thôn như sau:

	Ratio	SE	CV
Cả nước	0.0241752	0.0004209	1.74104
Thành thị	0.0491259	0.0009899	2.01503
Nông thôn	0.0156844	0.0004489	2.86208

Giá trị hệ số biến sai thấp có nghĩa rằng ước tính chính xác. Như vậy, chúng ta thấy rằng ước tính tỷ lệ thất nghiệp của cả nước chính xác hơn so với ước tính tỷ lệ thất nghiệp khu vực thành thị và nông thôn, và ước tính tỷ lệ thất nghiệp khu vực thành thị chính xác hơn so với ước tính tỷ lệ thất nghiệp của khu vực nông thôn. Điều này cũng phù hợp với việc phân bổ mẫu trong điều tra: quy mô mẫu hộ điều tra của cả nước là 100.680, của khu vực thành thị là 61.800 so với của khu vực nông thôn là 38.880.

(tiếp theo trang 23)

ĐÁNH GIÁ CHẤT LƯỢNG SỐ LIỆU ĐIỀU TRA THỰC TRẠNG VIỆC LÀM... (tiếp theo trang 18)

Qua phân tích và yêu cầu đánh giá chất lượng trên cho thấy để nâng cao chất lượng của cuộc điều tra cần hạn chế tối thiểu tổng sai số điều tra (= sai số chọn mẫu + sai số phi chọn mẫu). Muốn vậy, trước hết cần xác định cỡ mẫu cần thiết với mức yêu cầu về độ tin cậy của các ước tính để đáp ứng yêu cầu trên. Thứ hai, cần thiết kế quá trình đảm bảo chất lượng để xác định các hoạt động cụ thể cho việc giám sát thực hiện điều tra. Thứ ba, thực hiện việc hiệu đính bằng tay cũng như trên máy để kiểm tra, sửa những sai sót và loại trừ các nhân tố không cần thiết cũng như nhằm đạt được tính thống nhất của số liệu và tính thống nhất giữa các bảng số liệu.

Có thể nói, điều tra thực trạng việc làm và

thất nghiệp đã cung cấp một nguồn dữ liệu thông tin toàn diện về lực lượng lao động, việc làm, thiếu việc làm và thất nghiệp. Tuy nhiên cần phải tiếp tục nâng cao chất lượng và độ chính xác của các số liệu hiện nay. Với mục đích đó, việc tăng cường công tác tổ chức điều tra là một vấn đề đáng quan tâm. Việc tăng cường các phương tiện điều tra, cũng như cải thiện công tác thu thập và quản lý thông tin là những vấn đề cần được xem xét. Những cải tiến này cần phải được lập thành kế hoạch, và sắp xếp để đẩy mạnh công tác tổ chức thực hiện các cuộc điều tra trong những năm tới nhằm đảm bảo đạt được các mục tiêu, cũng như sử dụng hiệu quả các nguồn lực và phổ biến kịp thời kết quả điều tra ■