

Chung quanh vấn đề phương pháp xã hội học

Trong đợt hội thảo khoa học kỷ niệm 10 năm thành lập Viện Xã hội học tổ chức tại Viện, vấn đề phương pháp xã hội học được đặc biệt lưu ý và có nhiều tranh luận bổ ích.

Ở mục Diễn đàn Xã hội học kỳ này, chúng tôi giới thiệu ý kiến của một số cán bộ khoa học của Viện Xã hội học trình bày trong hội thảo khoa học nói trên.

TCXHH

Các cơ sở để lựa chọn phương pháp nghiên cứu

TÔN THIÊN CHIÊU

Cùng với sự phát triển của lý thuyết xã hội học, phương pháp xã hội học cũng không ngừng phát triển. Đến nay, phương pháp khảo sát Xã hội học rất phong phú, đa dạng và ngày càng hoàn thiện. Có rất nhiều tác giả nghiên cứu về phương pháp và đưa ra các quan điểm của mình về các phương pháp khác nhau được sử dụng trong điều tra xã hội học. Một câu hỏi đặt ra là trong điều tra xã hội học thì sử dụng hệ thống phương pháp nào là có hiệu quả nhất? Hiệu quả ở đây được hiểu là đem lại thông tin phản ánh đúng bản chất của quá trình, hiện tượng xã hội được nghiên cứu. Trả lời câu hỏi này không phải là đơn giản, bởi vì mỗi loại phương pháp đều có những ưu điểm và nhược điểm của nó. Các nhà xã hội học thực chứng thường chú trọng đến các phương pháp định lượng hơn. Còn các nhà xã hội học phản thực chứng lại chú ý đến các phương pháp định tính, cho rằng phương pháp định lượng không hiệu quả. Mỗi một trường phái như vậy đều có những quan điểm riêng của mình để bênh vực cho các phương pháp mà họ sử dụng.

Ở Việt Nam, thực tiễn hoạt động nghiên cứu xã hội học trong những năm qua cũng diễn ra tình hình như vậy. Mặt khác do xã hội học chỉ mới phát triển và được áp dụng trong

thời gian ngắn, cộng với những người làm công tác xã hội học không được đào tạo chính quy, cơ bản làm cho quan niệm về phương pháp xã hội học càng khác nhau. Mỗi người tiếp thu một trường phái xã hội học khác nhau, nên người thì đề cao một số phương pháp này, người khác lại chú trọng các phương pháp khác. Điều đó làm nảy sinh việc đề cao phương pháp này, coi phương pháp khác là không hiệu quả và còn dẫn đến nghi ngờ kết quả của người khác nếu người đó sử dụng khác phương pháp mà mình chưa hiểu rõ.

Trong bài viết này, chúng tôi không đề cập đến và cung chưa đủ khả năng để phân tích phương pháp nào là ưu việt hơn phương pháp nào, mà chỉ đưa ra một số cơ sở lý luận và thực tiễn cho sự lựa chọn phương pháp đối với từng cuộc điều tra. Trên cơ sở đó để mỗi người cân nhắc, lựa chọn phương pháp thích hợp cho cuộc điều tra phù hợp với nội dung mục đích và khả năng của mình.

I. NHỮNG YẾU TỐ ẢNH HƯỞNG ĐẾN CHẤT LƯỢNG THÔNG TIN TRONG ĐIỀU TRA THỰC NGHIỆM.

Mục đích của bất kỳ cuộc điều tra xã hội học nào cũng nhằm cung cấp cho chúng ta những thông tin cần thiết để phân tích, lý giải hiện tượng, quá trình xã hội mà ta đang nghiên cứu. Để phân tích và lý giải đúng đắn, thông tin thu được phải phản ánh đúng bản chất của hiện tượng, quá trình xã hội. Muốn vậy thông tin thu được phải mang tính đúng đắn (chân thực), khách quan.

Tính đúng đắn, khách quan của thông tin không chỉ phụ thuộc vào phương pháp nghiên cứu, mà còn phụ thuộc vào nhiều yếu tố khác như: hướng tiếp cận, mô hình lý luận, việc thao tác các chỉ báo (chọn biến số), cách thức tổ chức tiến hành thu thập thông tin, xử lý nó. Mỗi một yếu tố trên nếu không được chuẩn bị kỹ lưỡng, nghiêm túc đều ít nhiều tác động đến chất lượng thông tin làm cho thông tin không phản ánh đúng bản chất vấn đề nghiên cứu.

Đúng đầu trong các yếu tố ảnh hưởng thì phương pháp luận nghiên cứu ảnh hưởng mạnh nhất đến chất lượng thông tin. Bởi vì phương pháp luận không chỉ ảnh hưởng từ đầu mà còn ảnh hưởng đến các bước sau.

Phương pháp luận ở đây được chúng tôi dùng với ý nghĩa: chọn hướng tiếp cận, mô hình lý luận và xác định các chỉ báo (biến số).

Mỗi một quá trình, hiện tượng xã hội tồn tại và phát triển tự nhiên, nghĩa là nó tuân những quy luật chung và có quy luật đặc thù nội tại của nó. Song biểu hiện bên ngoài của nó lại rất đa dạng tùy từng lúc, từng địa điểm. Chính vì vậy mà việc nhận thức chúng cũng không phải mọi người đều giống nhau. Mỗi một quan điểm có một nhận thức khác nhau về vấn đề được nghiên cứu. Song hiện tượng, quá trình xã hội chỉ có một bản chất. Chính vì vậy, khi đó ra các quy tắc và phương pháp, E.Durkheim đã đặt nguyên tắc: "phải xem các sự kiện Xã hội như sự vật. Sự kiện xã hội là những "sự kiện" trực tiếp của khoa học, còn các nhận thức (lý luận) mà chúng ta có được đều do nhận thức từ sự kiện đó. Vì vậy nếu chúng ta không chọn đúng hướng tiếp cận sẽ dẫn đến nhận thức sai từ đầu bản chất của vấn đề Ví dụ: khi nghiên cứu phân tầng xã hội chúng ta có rất nhiều quan điểm lý luận (hướng tiếp cận) khác nhau. Mỗi hướng tiếp cận đó đều được xây dựng trên nền tảng của các xã hội khác nhau, ở mỗi thời điểm khác nhau. Vậy áp dụng quan điểm nào cho phù hợp với xã hội Việt Nam. điều này cần phải được cân nhắc.

Khi đã có hướng tiếp cận (quan điểm lý luận) rồi việc tiếp theo là xây dựng các mô hình lý luận (khung lý thuyết.) để thu thập thông tin cần thiết. Khung lý thuyết không chỉ là các căn cứ để thu thập mà còn dùng nó để phân tích thông tin sau này. Khi khung lý thuyết sau thì chẳng những thông tin thu sai và khi dùng nó để phân tích sẽ dẫn đến các

kết luận sai lầm. Thực tế nghiên cứu của chúng ta cũng đã có trường hợp như vậy. Một yếu tố khác trong phương pháp luận nghiên cứu ảnh hưởng đến chất lượng thông tin đó là việc lựa chọn các chỉ báo (biến số). Chẳng hạn khi nghiên cứu sự phân tầng xã hội, thì lấy chỉ báo nào để phân loại các gia đình thành các nhóm mức sống giàu nghèo khác nhau? "Lấy thu nhập, nhà ở, trang thiết bị gia đình hay chỉ tiêu làm chỉ báo chính? ở các nước công nghiệp người ta chỉ sử dụng chỉ báo thu nhập song ở Việt Nam, thực tiễn nếu chỉ lấy chỉ mỗi một chỉ báo này là chưa đủ mà còn dẫn đến sai lầm khi đánh giá mức sống của tầng lớp nhân dân.

Những điều đã nêu ở trên cho ta thấy rằng một khi cơ sở phương pháp luận của chúng ta đã sai thì dù có phương pháp đúng, cách tổ chức hoàn thiện... đều đưa đến hậu quả là thông tin không phản ánh đúng bản chất của vấn đề.

Phương pháp nghiên cứu cũng ảnh hưởng đến chất lượng thông tin. Chọn phương pháp không đúng còn dẫn đến việc tốn kém tiền của và thời gian. Chẳng hạn, nếu mục đích của cuộc nghiên cứu chỉ là phát hiện vấn đề, đề ra các giả thuyết thì rõ ràng ta không cần phải dùng các phương pháp nghiên cứu định lượng, có quy mô lớn, mà chỉ cần các cuộc phỏng vấn sâu có quy mô ít. Ngược lại, khi muốn chứng minh hoặc khẳng định một giả thuyết nào đó, thì không thể tiến hành các cuộc nghiên cứu có quy mô nhỏ, trên chỉ một hai địa bàn đặc biệt, bởi vì thông tin thực nghiệm luôn luôn chịu sự chi phối của không gian và thời gian thu thập.

Yếu tố tổ chức cũng ảnh hưởng đến chất lượng thông tin. Yếu tố này bao gồm: lựa chọn, tập huấn điều tra viên, chọn thời điểm điều tra... Chẳng hạn đối với việc chọn điều tra viên, mỗi cách thu thập thông tin có một yêu cầu riêng. Điều tra bằng bảng hỏi thì không cần đội ngũ điều tra viên lão luyện, giàu kinh nghiệm, có hiểu biết bằng khi phỏng vấn. Ngay trong phỏng vấn thì tiêu chuẩn của điều tra viên cho phỏng vấn tự do (không cấu trúc) cũng phải khác với phỏng vấn tiêu chuẩn (phỏng vấn cấu trúc). Vai trò của điều tra viên trong việc thu thập thông tin sơ cấp là cực kỳ quan trọng. Điều tra viên là cầu nối giữa người nghiên cứu, chủ đề nghiên cứu với những người cung cấp thông tin. Điều tra viên có trách nhiệm, có kinh nghiệm và phù hợp sẽ hạn chế được những thông tin không trung thực do nguồn cung cấp đưa ra. Nói đến điều tra viên cũng là nói đến vấn đề tài chính. Phải có một khoản tài chính nào đó mới khuyến khích được điều tra viên hăng hái, có trách nhiệm.

Với các cuộc điều tra định lượng việc xử lý các số liệu cũng có ảnh hưởng nhất định đến chất lượng thông tin. Người xử lý, nếu không hiểu vấn đề nghiên cứu cũng có thể đưa cho chúng ta các kết quả sai lầm.

II. CÁC PHƯƠNG PHÁP THU THẬP THÔNG TIN VÀ DẠNG NGHIÊN CỨU

Bấy lâu nay chúng ta thường có những lẫn lộn đồng nhất giữa dạng nghiên cứu và phương pháp nghiên cứu, giữa phương pháp nghiên cứu và phương pháp thu thập thông tin. Phương pháp thu thập thông tin nêu lên cách thức ta thu thập thông tin nào. Phương pháp nghiên cứu không chỉ nêu lên mỗi một cách thu thập thông tin mà còn bao hàm cả phương pháp chọn mẫu. Còn dạng nghiên cứu là hình thức chúng ta tiến hành một cuộc nghiên cứu không chỉ sử dụng một phương pháp, mà có thể hai, ba phương pháp.

Sở dĩ có sự đồng nhất đó là do ứng với mỗi dạng nghiên cứu đều kèm với một phương pháp nghiên cứu, và một phương pháp thu thập thông tin có hiệu quả. Chẳng hạn phương pháp nghiên cứu sâu, thì hiệu quả nhất là phỏng vấn sâu. Còn dạng nghiên cứu thâm nhập

thực tế có thể sử dụng nhiều phương pháp song tốt nhất là nghiên cứu sâu và quan sát.

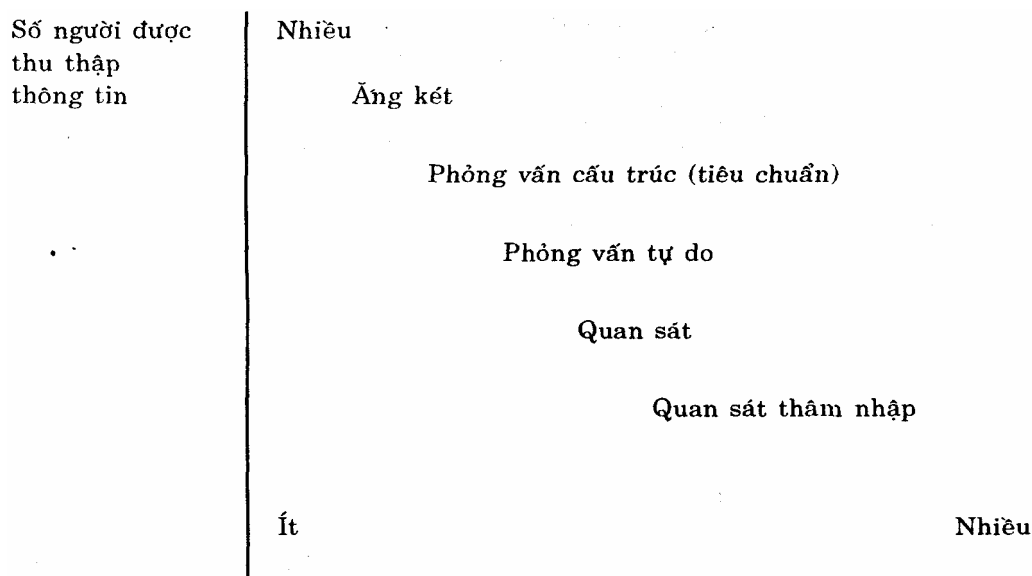
1. Các phương pháp thu thập thông tin

Ngày nay để tiến hành thu thập thông tin sơ cấp, xã hội học có nhiều phương pháp khác nhau. Song nếu theo tầm quan trọng của vai trò, trách nhiệm và khả năng của điều tra viên mà chúng ta có thể quy lại thành 3 dạng chính: ăng ket - Phỏng vấn - Quan sát. Trong mỗi dạng đều có những biến dạng của nó. Xếp theo vai trò của điều tra viên từ thấp đến cao là:

- Ăng ket
- Phỏng vấn tiêu chuẩn (có bảng hỏi, được cấu trúc chặt chẽ)
- Phỏng vấn tự do (phỏng vấn sâu, có thể có bảng hỏi hoặc không, song cấu trúc không chặt, phỏng vấn nhóm).
- Quan sát
- Quan sát thâm nhập.

Mỗi một loại phương pháp thu thập thông tin này đều có những ưu điểm và nhược điểm riêng, tùy thuộc vào từng phương pháp nghiên cứu, dạng nghiên cứu. Chính vì vậy mà mỗi một phương pháp thu thập được dùng với một dạng nghiên cứu cụ thể. Ví dụ ăng ket thường được dùng trong phương pháp định lượng và dạng nghiên cứu định lượng. Trong khi đó phỏng vấn sâu thường lại được dùng trong nghiên cứu sâu.

Ở đây chúng tôi không đi vào phân tích các ưu, khuyết điểm của từng loại, cũng như các điều kiện khi sử dụng chúng. Chúng tôi chỉ đưa ra biểu đồ minh họa mối liên hệ của các phương pháp thu thập thông tin và tính chất của các kết luận rút ra qua những thông tin thu được chúng, cũng như mức độ người được phỏng vấn. Qua biểu đồ này cũng giúp cho ta hiểu thêm chọn phương pháp nào cho phù hợp với mục đích, nội dung của từng cuộc nghiên cứu, và tính chất của các kết luận đưa ra sau cuộc nghiên cứu.



Tính chủ quan của người nghiên cứu khi rút ra các kết luận.

Còn một phương pháp thu thập thông tin nữa là sử dụng tài liệu có sẵn. Song ở đây không phải là thông tin sơ cấp, mà là thông tin thứ cấp. Thông tin thứ cấp là thông tin đã được thu thập từ trước theo một mục đích nào đó của người lập ra thông tin. Khi sử dụng lại thông tin này người nghiên cứu cần chú ý đến mục đích và phương pháp thu thập thông tin của họ, để sử dụng có hiệu quả hơn.

2- Dạng nghiên cứu

Việc phân biệt các dạng nghiên cứu có nhiều cách phân chia khác nhau. Có người phân chia dựa theo phương thức thu thập thông tin, song cũng có người phân chia dựa trên tiêu chuẩn phương pháp rút ra kết luận. Mặc dù có nhiều cách phân chia như vậy song cũng có những điểm trùng nhau.

Theo cách phân chia theo phương pháp thu thập thông tin ta có các dạng nghiên cứu sau:

- Thử nghiệm xã hội (tạo tình huống)
- Thâm nhập thực tế
- Sử dụng tài liệu có sẵn
- Điều tra chọn mẫu

Tính khái quát của các kết luận phát triển theo bậc thang đi từ thử nghiệm xã hội đến điều tra chọn mẫu. Nghĩa là các kết quả thu được trong thử nghiệm xã hội (tạo tình huống) mang tính chất cục bộ, riêng lẻ, đơn nhất.

Như chúng tôi đã nêu ở trên, mỗi dạng nghiên cứu có thể sử dụng nhiều phương pháp thu thập thông tin, song mỗi dạng có một hai phương pháp chủ đạo, thích hợp. Chính vì vậy mà có thể gọi là phương pháp nghiên cứu. Ví dụ dạng nghiên cứu thâm nhập thực tế có người còn gọi là phương pháp thâm nhập thực tế. Bảng sau cho ta được mối liên hệ giữa dạng nghiên cứu và phương pháp thu thập thông tin.

Dạng nghiên cứu	Phương pháp thu thập thông tin	
	Chủ yếu	Bổ trợ
	Quan sát	Phỏng vấn sâu
- Tạo tình huống		
- Thâm nhập thực tế	Phỏng vấn sâu quan sát	Nghe báo cáo (Tọa đàm...)
- Điều tra chọn mẫu	-Ăng két... - Phỏng vấn tiêu chuẩn	- Quan sát - Phỏng vấn tự do - Tọa đàm...
- Sử dụng tài liệu có sẵn	Tài liệu có sẵn	

Về dạng nghiên cứu ta còn một cách nữa dựa trên phương pháp phân tích số liệu sơ cấp là: nghiên cứu định lượng và nghiên cứu định tính. Nghiên cứu định lượng là dạng nghiên cứu dựa vào phương pháp phân tích số liệu định lượng để rút ra kết luận. Còn nghiên cứu định tính, để rút ra kết luận người nghiên cứu phải sử dụng tư duy lý luận. Việc phân chia ranh giới giữa định lượng và định tính là tương đối, bởi vì để định lượng được cần phải có

các phân tích định tính trước đã. Nếu định lượng mà không được định tính từ trước dẫn đến các kết quả thiếu thuyết phục. Nghiên cứu định lượng là bước phát triển kế tiếp của nghiên cứu định tính.

Nếu như nghiên cứu định tính nhằm mục đích phát hiện vấn đề, xây dựng giả thuyết, giải thích... thì nghiên cứu định lượng hướng đến mức độ của vấn đề, kiểm tra giả thuyết đúng hay sai... Ví dụ: nghiên cứu định tính cho ta biết hiện nay ở xã hội ta đang có sự phân hóa xã hội thành giàu nghèo. Song nếu chỉ dừng lại ở nghiên cứu định tính (nghĩa là không dùng các phương pháp định lượng để phân tích số liệu) thì sẽ không trả lời cho ta được câu hỏi: Vậy giàu là bao nhiêu và nghèo là bao nhiêu? Mức độ chênh lệch giữa hai nhóm này là thế nào? Nghiên cứu định lượng còn có thể trả lời được sâu hơn một số câu hỏi mà định tính không trả lời được. Chẳng hạn: Nghiên cứu định tính xác nhận có một sự ảnh hưởng của học vấn của bố mẹ và mức sống của gia đình đến việc hướng nghiệp cho con. Song, học vấn ảnh hưởng mạnh hay mức sống mạnh hơn, nghiên cứu định tính chỉ cho trả lời cho từng trường hợp (gia đình) cụ thể. Nghiên cứu định lượng, có thể trả lời được câu hỏi đó trong đa số dân cư. Khi nghiên cứu định tính phương pháp thu thập thông tin phong phú hơn, còn nghiên cứu định lượng chỉ có ăng kết và phỏng vấn tiêu chuẩn hoá.

Nghiên cứu định lượng có yêu cầu và một tổ chức chặt chẽ khó khăn hơn so với nghiên cứu định tính. Chẳng hạn nghiên cứu định tính không quan tâm đến việc chọn mẫu, còn nghiên cứu định lượng chọn mẫu lại rất quan trọng. Số lượng người được nghiên cứu trong định tính cũng không có yêu cầu bắt buộc. Trong khi đó, điều tra bao nhiêu người thì đạt được yêu cầu (sai số cho phép) cũng được đặt ra trong nghiên cứu định lượng. Nếu số người quá ít các con số nêu ra sẽ mất hết ý nghĩa.

Mối liên hệ giữa hai dạng phân chia được thể hiện như sau:

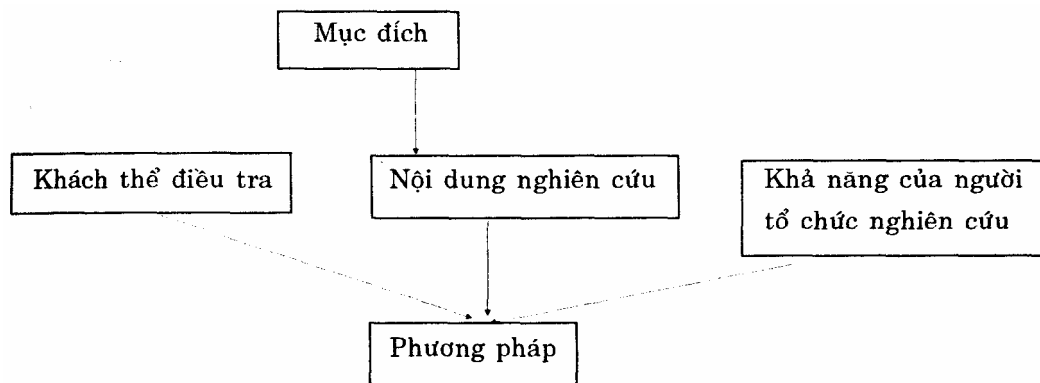
- | | |
|--------------|---------------------------|
| - Định lượng | + Điều tra chọn mẫu |
| | + Sử dụng tài liệu có sẵn |
| - Định tính | + Case study |
| | + Tạo tình huống |

Qua sơ đồ trên, chúng ta thấy các cuộc điều tra chọn mẫu là nghiên cứu định lượng, còn tạo tình huống là nghiên cứu định tính. Còn sử dụng tài liệu có sẵn và Case study là nằm giáp ranh của cả hai dạng. Đối với nghiên cứu trường hợp, trước đây khái niệm này thường chỉ dùng để chỉ cá nhân, song cùng với thời gian nó có thể rộng cho cả một làng và có thể sử dụng các phương pháp định lượng trong trường hợp đó. Tuy vậy Case study vẫn chủ yếu là nghiên cứu định tính vì nó không thờ rộng, phổ quát kết quả rộng ra được. Trường hợp phân tích tài liệu có sẵn cũng vậy. Tùy theo sử dụng cách phân tích nào mà ta có dạng nghiên cứu định tính hay định lượng.

III- CÁC CƠ SỞ ĐỂ LỰA CHỌN PHƯƠNG PHÁP NGHIÊN CỨU

Ở trên, chúng ta đã thấy được các yếu tố ảnh hưởng đến chất lượng thông tin và đặc thù của từng phương pháp nghiên cứu. Vậy khi nhận một đề tài nghiên cứu, chúng ta căn cứ vào đầu để lựa chọn phương pháp nghiên cứu cho thích hợp. Sẽ là vội vàng nếu chúng ta trả lời ngay phương pháp này hay phương pháp kia khi chưa cân nhắc đầy đủ các dữ liệu cấu thành cho một cuộc nghiên cứu.

Bảng sơ đồ sau đây, cho phép ta cân nhắc đầy đủ các yếu tố khi lựa chọn phương pháp điều tra.



Yếu tố cơ bản quyết định đến phương pháp nghiên cứu là nội dung, mục đích nghiên cứu, song nếu chỉ căn cứ vào các yếu tố này mà không cân nhắc đến hai yếu tố khác: Khả năng của người tổ chức nghiên cứu, và khách thể điều tra thì việc lựa chọn phương pháp cũng không đảm bảo mang lại cho chúng ta thông tin đảm bảo chất lượng.

Khi nói đến mục đích, nội dung của cuộc nghiên cứu làm căn cứ cho việc lựa chọn phương pháp những điều cần chú trọng là: Mức độ khái quát và độ chính xác, mức độ sâu rộng của nội dung nghiên cứu.

Mức độ khái quát của các kết quả thu được cho phép ta xác định cuộc điều tra này phải là định tính hay định lượng. Khi đã nghiên cứu định lượng thì độ chính xác là bao nhiêu? Điều đó cho ta hình dung cách chọn mẫu như thế nào cho đại diện, và dung lượng mẫu là bao nhiêu.

Mức độ sâu, rộng của nội dung giúp ta cân nhắc nội dung nào cần thu thập thông tin bằng cách khác để tránh làm cho một phương pháp thu thập thông tin quá tải, không hiệu quả.

Yếu tố khả năng của người nghiên cứu thể hiện ở bình diện: Trình độ của người nghiên cứu và tài chính có được cũng như thời gian. Chẳng hạn với một cuộc nghiên cứu định lượng ngoài những trình độ lý luận vốn có của người nghiên cứu, thì kỹ năng soạn thảo bảng hỏi (cách đặt câu hỏi, bố trí logic các câu hỏi), nhất là trong phương pháp ăng két, cũng chiếm một vị trí đặc biệt. Vấn đề tài chính cũng quyết định không nhỏ tới việc lựa chọn phương pháp điều tra. Nếu tài chính eo hẹp sẽ hạn chế việc tuyển chọn điều tra viên phù hợp, có trách nhiệm, kinh nghiệm và đến dung lượng mẫu, ảnh hưởng đến mức chính xác và khái quát của các kết quả rút ra hoặc chất lượng thông tin.

Một yếu tố nữa, mà khi lựa chọn phương pháp chúng ta cần quan tâm đó là: chính bản thân người được điều tra. Thực ra chất lượng thông tin hoàn toàn vào người cung cấp thông tin, song việc lựa chọn phương pháp thích hợp, đặt các câu hỏi phù hợp, chính là hạn chế bớt ảnh hưởng của người cung cấp thông tin thêm trung thực.

Thực chất của việc lựa chọn phương pháp điều tra là giải bài toán cân đối các yếu tố trên, làm sao chọn được phương pháp đáp ứng gần đủ các yêu cầu của các yếu tố. Như vậy, một phương pháp được hỏi là tối ưu, có hiệu quả là phương pháp đáp ứng được các yêu cầu của một cuộc nghiên cứu đề ra. Việc đề cao từng phương pháp một cách chung chung mà không căn cứ vào từng cuộc điều tra theo mục đích yêu cầu của nó là không thỏa đáng. Mỗi một phương pháp ưu việt đối với dạng nghiên cứu này mà không ưu việt chỉ với dạng nghiên cứu khác, mà có thể dẫn đến những kết luận sai lầm.

Tổng quan về các nghiên cứu chọn mẫu qua các hoạt động của Phòng Xã hội học dân số và gia đình

PHẠM BÍCH

Các cuộc nghiên cứu chọn mẫu? Mẫu gì vậy? chúng tôi chưa làm bao giờ. Nghiên cứu xã hội học? Chúng tôi thường xuyên làm. Hơn thế, các bạn có thể thấy số liệu ở đây trong số những người được hỏi có tỷ lệ phần trăm... những người trả lời nói rằng...

Các câu nói trên, có thể nhiều người cho là quá sự thật, nhưng có lẽ cũng không xa sự thật là bao trong bối cảnh của nền xã hội học Việt Nam hiện nay. Bên ngoài Viện Xã hội học, phần lớn các cơ quan bạn vẫn tiến hành các nghiên cứu Xã hội học. Tuy nhiên, theo cách hiểu thông thường thì nghiên cứu chọn mẫu đã được đồng nhất với nghiên cứu xã hội học, và ngoài ra không còn có phương pháp gì khác hơn nữa. Quá hơn, nếu không có các con số phần trăm thì cuộc nghiên cứu của bạn sẽ là không có giá trị khoa học và, thậm chí, là tư biện.

Như vậy, phương pháp xã hội học nghiên cứu chọn mẫu với nền tảng tiếp cận nguyên thủy là chủ nghĩa thực chứng trong khoa học, trường phái thịnh hành vào nửa cuối thế kỷ trước và khoảng ba thập kỷ đầu của thế kỷ XX, cho đến nay vẫn là phương pháp thịnh hành nhất, nếu không phải là duy nhất hiện đang được áp dụng ở Việt Nam. Từ ngày E. Durkheim triển khai nghiên cứu hiện tượng tử thông qua việc xem xét số liệu thống kê của các nước khác nhau cùng với sự tiến triển của kỹ thuật thống kê, phương pháp nghiên cứu chọn mẫu đã đạt được những thành quả ngày càng lớn lao. Việc xem xét sự phát triển của các nghiên cứu chọn mẫu đã được tiến hành tại Viện Xã hội học trong mười năm qua sẽ cung cấp cho chúng ta một ý niệm về sự phát triển của phương pháp này nói riêng cũng như nghiên cứu xã hội học nói chung của Viện. Trong bài viết này chúng tôi chọn Case study là các nghiên cứu chọn mẫu của Phòng Xã hội học dân số và gia đình.

Nhìn chung lại, các nghiên cứu chọn mẫu của Phòng Xã hội học dân số và gia đình đã trải qua những giai đoạn sau đây.

I. Giai đoạn thử nghiệm

Giai đoạn thử nghiệm được tiến hành từ khoảng những ngày đầu thành lập Viện cho đến khoảng giữa năm 1984. Giai đoạn này có ba cuộc nghiên cứu quan trọng nổi bật: nghiên cứu Hải Vân, nghiên cứu Vĩnh Phú theo yêu cầu của Hội Liên hiệp phụ nữ và nghiên cứu Quyết Tiến. Nếu cuộc nghiên cứu Hải Vân là cuộc nghiên cứu nhập môn cho các nghiên cứu xã hội học hiện đại đối với nhiều chương trình nghiên cứu trong Viện thì hai cuộc nghiên cứu còn lại đúng là những nghiên cứu thử nghiệm cho nghiên cứu gia đình và nghiên cứu dân số.

Ngoài ra còn có thể kể đến những cuộc nghiên cứu khác như nghiên cứu Hải Thanh (11/1984, Hải Hậu, Hà Nam Ninh), nghiên cứu Bình Đà (10/1984, Thăng Bình, Quảng Nam Đà Nẵng), nghiên cứu di dân Đông - Tây Nam Bộ (7/1984)...

Đặc trưng cơ bản của các cuộc nghiên cứu thời kỳ này (không kể cuộc nghiên cứu Hải Vân) là sự vắng bóng một tiếp cận lý thuyết rõ ràng và điều đương nhiên là sự không có một định hướng chính xác về phương pháp, cái được thể hiện qua hệ biến số và hệ chỉ báo.

Bảng hỏi thường được ghép nối từ nhiều chuyên gia khác nhau (nghiên cứu Vĩnh Phú) và yêu cầu xử lý thường chỉ bắt đầu được nghĩ tới sau khi đã thu thập số liệu điền dã về. Chọn mẫu tại nông thôn thường được tiến hành căn cứ vào sổ hộ tịch xã. Độ ổn định và độ xác thực (reliability và validity) của câu hỏi chưa được nghĩ tới. Có thể nói các nghiên cứu này có tính chất thử nghiệm - sai lầm - điều chỉnh.

II. Giai đoạn lựa chọn lý thuyết (1985 - 1989)

Tầm quan trọng của lý thuyết được thể hiện rõ nét nhất trong quá trình triển khai nghiên cứu theo đề tài của Viện Xã hội học nằm trong chương trình nghiên cứu cấp nhà nước Tây Nguyên II. Lý thuyết được lựa chọn ban đầu cho cuộc nghiên cứu này là lý thuyết hút - đẩy trong quá trình di dân của E.Lee. Lý thuyết này làm việc tốt trong cuộc nghiên cứu di dân tại các khu vực miền Đông - Tây Nam Bộ nhưng không hoạt động tại Tây Nguyên do các yêu cầu của cuộc nghiên cứu, dù nghiên cứu qua người di dân, nhưng lại không chỉ là vấn đề di dân. Sau hai cuộc điền dã không thành công tại Tây Nguyên, 1985/1986, một khung lý thuyết mới được xây dựng và chỉ sau đó các kết quả điền dã mới được xử lý theo một khung phản ánh được yêu cầu nghiên cứu.

Một cuộc nghiên cứu khác đáng được lưu ý tới trong thời kỳ này là nghiên cứu. Dân số và Sức khỏe/1988 với xấp xỉ 5000 mẫu trên địa bàn toàn quốc. Khung lý thuyết được truyền vào Việt Nam không đầy đủ và sai lệch nhiều do trình độ của những người tiến hành, do vậy các câu hỏi không thể hiện đầy đủ hệ biến số độc lập và phụ thuộc theo nguyên bản quốc tế Tuy nhiên cách chọn mẫu phân tổ ngẫu nhiên của các chuyên gia quốc tế đã cứu vãn nhiều cho kết quả nghiên cứu. Kết quả thu được trên thực tế, tuy chưa được làm sạch, phản ánh phần nào những gì đã diễn ra ở Việt Nam trong lĩnh vực dân số và sức khỏe.

Kinh nghiệm của thời kỳ này đã buộc các nghiên cứu chọn mẫu phải tuân thủ một lý thuyết với khung nghiên cứu rất rõ ràng.

III. Giai đoạn phương pháp thể hiện lý thuyết (1990 - 10/1993)

Giai đoạn này được thể hiện bằng một loạt cuộc nghiên cứu đặc sắc và sự hoàn thiện căn bản kiến thức về phương pháp nghiên cứu chọn mẫu. Điều này diễn ra cùng với quá trình đổi mới khi khả năng tiếp cận với bên ngoài gia tăng và các tài liệu kinh nghiệm nghiên cứu thu thập được qua các cuộc trao đổi với các học giả bên ngoài được nâng cao. Trước hết phải kể đến ở đây là cuộc nghiên cứu "sự biến đổi gia đình và mức sinh" (FFS li990) tại ba khu vực Bắc, Trung, Nam với một khung lý thuyết chặt chẽ và các biến số cũng như hệ thống chỉ báo được xây dựng nghiêm túc trước khi điền dã. Mẫu được chọn theo phương pháp ngẫu nhiên có chủ định căn cứ vào sổ hộ tịch. Tuy nhiên, sự yếu kém về phương tiện xử lý đã không cho phép khai thác các điểm mạnh của bảng hỏi.

Cuộc nghiên cứu thử nghiệm "*Lịch sử đời sống và Động thái dân số*" (1991) hợp tác với giáo sư C. Hirschman đem lại kinh nghiệm về cách tiếp cận lịch sử đời sống, hệ các biến số và chỉ báo và đặc biệt là bài học về quá trình làm sạch hết sức tỉ mỉ các số liệu thu được trước khi đưa vào xử lý.

Các cuộc nghiên cứu tham gia vào đề tài KX - 04 - 02 ở Hà Nội, Hải Phòng, Quảng Nam - Đà Nẵng, thành phố Hồ Chí Minh, Cần Thơ đặt ra vấn đề phải có sự gia trọng để có thể so sánh được các số liệu của các nhóm khác nhau, đặc biệt là thu nhập của hộ gia đình tại các khu vực khác nhau và địa bàn nông thôn, đô thị khác nhau. Điều này dẫn đến khả năng so sánh số liệu trong tương lai tạo nên hệ thống số liệu chung có thể chấp nhận được ở quy mô toàn quốc.

Cuối cùng, cuộc nghiên cứu *Kiến thức, thái độ và việc thực hiện các biện pháp kế hoạch hóa gia đình* (KAP/93) đánh dấu sự hoàn thiện về ý thuyết, hệ biến số và chỉ báo, kỹ thuật máy tính khai thác thông tin thu được. Có thể nói sau cuộc nghiên cứu này chúng tôi đã nắm tương đối vững phương pháp nghiên cứu chọn mẫu.

IV. Giai đoạn đi sâu và đặc trưng ý nghĩa thống kê (10/1993)

Cuộc nghiên cứu KAP/93 đã đặt ra một vấn đề gay gắt là làm thế nào tiến hành chọn mẫu tốt trên quy mô toàn quốc. Đồng thời cuộc hợp tác nghiên cứu giữa Việt Nam và Indonesia thông qua xử lý lại số liệu của *Nghiên cứu Dân số và Sức khỏe 188* (VN/DHS/88) đặt ra một vấn đề khác là ý nghĩa thống kê của các số liệu đưa ra. Nhưng vấn đề này đòi hỏi các nhà nghiên cứu phải triển khai lại việc nghiên cứu cơ bản lý thuyết thống kê và công nghệ áp dụng trong các cuộc nghiên cứu chọn mẫu. Đây là giai đoạn cuối cùng, theo ý kiến của chúng tôi, để nắm vững phương pháp nghiên cứu chọn mẫu xã hội học. Nét đặc trưng của Viện Xã hội học là một quá trình xã hội học vừa học vừa làm. Quá trình đó đòi hỏi một loạt các thử nghiệm - sai lầm - điều chỉnh với sự thể hiện rất rõ ràng trong các nghiên cứu chọn mẫu của chương trình dân số và gia đình. Vấn đề đặt ra ở đây là làm thế nào để có thể rút ngắn thời gian thử nghiệm ấy lại và làm thế nào để cho người nghiên cứu khi tiến hành nghiên cứu phải nắm vững phương pháp nghiên cứu của mình. Có lẽ, đó là nhiệm vụ hết sức quan trọng mà mười năm xây dựng Viện Xã hội học Việt Nam theo hướng hiện đại đặt ra cho chúng ta hôm nay.

Nghiên cứu trường hợp (Case study)

như một chiến lược nghiên cứu

THÀNH NHÂN

1. Case study trong các chiến lược nghiên cứu

Khi nào và tại sao cần tiến hành một Case study trên một chủ đề nào đó? Có nên cân nhắc sử dụng, chẳng hạn một nghiên cứu thử nghiệm, một survey hay một nghiên cứu lịch sử để thay thế không?

Mỗi một sự lựa chọn như vậy là một chiến lược nghiên cứu khác nhau. Mỗi chiến lược tại là một cách thức thu thập, phân tích những bằng chứng thực nghiệm khác nhau. Và mỗi cái đều có ưu và nhược điểm riêng.

Có một quan niệm không chính xác song khá phổ biến là: các chiến lược nghiên cứu khác nhau cần được sắp xếp theo thứ hạng. Chẳng hạn.

- Case study thích hợp cho giai đoạn thăm dò của 1 cuộc nghiên cứu.
- Survey và nghiên cứu lịch sử thích hợp cho giai đoạn mô tả.
- Còn thử nghiệm chỉ nên dùng trong các khảo sát giải thích và tìm các quan hệ nhân quả v v ...

Trên thực tế thì không hẳn như thế. Ở đây ranh giới giữa các chiến lược nghiên cứu

là không quá chặt chẽ. Mỗi chiến lược có thể được sử dụng cho cả 3 mục đích: thăm dò mô tả giải thích. Một quan niệm phù hợp hơn về việc sử dụng các chiến lược nghiên cứu có lẽ là quan niệm đa nguyên (pluralistic). Cần tránh những sai lầm quá thô thiển như: đáng lý sử dụng chiến lược này thì lại sử dụng chiến lược khác kém ưu thế hơn hẳn. Có một cách phân biệt tốt hơn Case study với các chiến lược nghiên cứu khác là so sánh 3 điều kiện tình huống như sau:

Các tình huống thích hợp cho những chiến lược nghiên cứu khác nhau

Chiến lược nghiên cứu	Hình thức của các câu hỏi (vấn đề) nghiên cứu	Đòi hỏi có sự kiểm soát đối sự kiện hành vi	Tập trung vào các sự kiện hiện tại
Thử nghiệm	Như thế nào tại sao?	Có	Có
Survey (nghiên cứu mẫu)	Ai, cái gì, ở đâu, bao nhiêu?	Không	Có
Phân tích tài liệu	Ai, cái gì, ở đâu	Không	Có/ Không
Thống kê lưu trữ	bao nhiêu?		
Lịch sử	Như thế nào, tại sao?	Không	Không
Case study	Như thế nào, tại sao?	Không	Có

Có hai nhóm câu hỏi chính:

Nhóm câu hỏi: "Ai, cái gì, ở đâu" tương đối thích hợp để được đặt trong các nghiên cứu thăm dò, mô tả. Mở rộng ra có thêm các câu hỏi "bao nhiêu" để định lượng hóa.

- Nhóm câu hỏi "Như thế nào, tại sao" rõ ràng là thích hợp cho các nghiên cứu giải thích, ở đây:

+ Các chiến lược nghiên cứu có thể được sử dụng linh hoạt riêng rẽ hoặc kết hợp (Ví dụ có các Case study trong một cuộc survey hoặc ngược lại).

+ Các chiến lược nghiên cứu là không loại trừ nhau mà bổ sung cho nhau.

Tuy nhiên, có những tình huống mà một chiến lược nghiên cứu nào đó là thuận lợi nhất. Đối với Case study đó là tình huống: khi mà các câu hỏi "Như thế nào, tại sao?" được đặt ra cho một loạt sự kiện xảy ra trong hiện tại mà nhà nghiên cứu có rất ít hoặc không có sự theo dõi và kiểm soát (hay hiểu biết) đối với chúng.

Ở đây chức năng thăm dò và giải thích là rất rõ. Hai câu hỏi "Như thế nào và tại sao" được nêu ra trong nghiên cứu là một ví dụ có tính kinh điển về tình huống sử dụng Case study.

2. Một định nghĩa về Case study

- Có rất nhiều định nghĩa

Đây là một định nghĩa mang tính “kỹ thuật” nhiều hơn. Case study là một khảo sát thực nghiệm mà:

- + Nghiên cứu một hiện tượng đang xảy ra trong bối cảnh đời sống hiện thực của nó, khi mà:
- + Những ranh giới giữa hiện tượng và bối cảnh là không rõ ràng,
- + Sử dụng nhiều nguồn thông tin, bằng chứng khác nhau.

Định nghĩa này không chỉ giúp ta hiểu về Case study mà còn giúp phân biệt nó với các chiến lược nghiên cứu khác.

Ví dụ

- Thử nghiệm: *Tách biệt* hoàn toàn hiện tượng - Bối cảnh .
- Lịch sử. Có chú ý tới tình huống phức tạp của Hiện tượng và Bối cảnh, *song trong quá khứ*
- Survey: cố gắng gắn với cả Hiện tượng lẫn Bối cảnh. Song khả năng của Survey nghiên cứu Bối cảnh là cực kỳ hạn chế. Người thiết kế Survey phải thường xuyên “chiến đấu” để hạn chế số lượng các biến số cần phân tích (hay số câu hỏi cần hỏi) và cố gắng cho về an toàn trong khuôn khổ dung lượng mẫu có thể nghiên cứu.

3. Case study trong các nghiên cứu đánh giá

Case study có vai trò nổi bật trong việc đánh giá, chẳng hạn một sự can thiệp nào đó một tác động của một chính sách, biện pháp cụ thể nào đó.

Có hướng áp dụng Case study trong lĩnh vực này

- a) Quan trọng nhất là *giải thích các quan hệ nhân quả trong sự can thiệp vào đời sống hiện thực* mà quá phức tạp để sử dụng các chiến lược Survey hoặc thử nghiệm:
- b) Mô tả bối cảnh đời sống - hiện thực trong đó sự can thiệp đã diễn ra.
- c) Case study đóng vai trò minh họa cho sự can thiệp, giúp vào việc đánh giá sự can thiệp này.
- d) Sử dụng Case study để thăm dò phát hiện những tình huống trong đó, sự can thiệp cần đánh giá chưa cho những hệ quả rõ ràng, xác định.

Tiếp cận chức năng trong nghiên cứu xã hội học

VŨ TUẤN HUY

Tiếp cận chức năng là một trong những mô hình nghiên cứu xã hội học mang tính lý thuyết, phương pháp luận hơn là những phương pháp cụ thể để thu thập thông tin xã hội học. Tính lý thuyết của nó biểu hiện ở chỗ nó trả lời câu hỏi: "Vấn đề xã hội nào sẽ được nghiên cứu và liên quan đến câu hỏi đó là những thông tin gì sẽ được thu thập?". Nó chỉ ra các chiều của sự kiện xã hội, ở những quy mô khác nhau. Do đó thể là toàn bộ xã hội như một tổng thể, hoặc một thiết chế, một nhóm. Vì vậy, tiếp cận chức năng là cái khung để thu thập thông tin xã hội học đồng thời là khung lý thuyết để giải thích các hiện tượng xã hội đó

Tiếp cận chức năng bắt đầu với ý tưởng là bất kỳ một hệ thống ổn định nào cũng bao gồm một số các bộ phận khác nhau, nhưng liên hệ qua lại với nhau và cùng hoạt động tạo nên sự cân bằng của toàn bộ hệ thống. Mỗi bộ phận tồn tại trong hệ thống không thể chỉ vì nó tồn tại. Nó phải có chức năng nào đó góp phần vào sự ổn định của hệ thống. Đó là các chức năng tích cực (positive function). Đương nhiên, các bộ phận khác nhau trong hệ thống có tầm quan trọng khác nhau về mặt chức năng (functional importance).

Thuyết chức năng của Talcott Parsons được phát triển vào cuối những năm 1930 với quan niệm rằng xã hội là một hệ thống có cấu trúc phức tạp với nhiều bộ phận liên quan với nhau ở nhiều cấp độ. Các cấp độ chủ yếu là 1) hữu cơ; 2) nhân cách; 3) cấu trúc; 4) văn hóa. Parsons cho rằng sự nhất trí cao giữa các thành viên của nhóm là lực lượng chủ yếu tạo nên sự ổn định của xã hội như một hệ thống nói chung. Các thành viên của xã hội học được xã hội hóa theo những chuẩn mực và giá trị bên trong của nhóm. Những hành vi lệch chuẩn tạo nên sự bất ổn định của nhóm được xem là những yếu tố bên ngoài.

Nhấn mạnh đến quan điểm này, thuyết chức năng của Parsons khó có thể giải thích những vấn đề như hành vi tội phạm, hiện tượng nghèo khổ và những hành vi lệch chuẩn khác vẫn tồn tại dai dẳng trong các xã hội. Và nếu như nó là một bộ phận trong cơ cấu xã hội thì, việc giải thích các chức năng của nó sẽ mâu thuẫn với cách giải thích theo quan điểm của Parsons.

Robert Merton đã đưa ra một quan điểm khác về chức năng. Ông cho rằng không phải mọi hành vi, ý tưởng của con người đều có chức năng tích cực như nó vẫn thường được xem là như vậy. Ông gọi hậu quả của những hành vi dẫn đến phá vỡ sự ổn định của hệ thống là các phản chức năng (dysfunction). Do đó, một hành vi, một mục đích có thể dẫn đến sự ổn định trong hệ thống này nhưng lại phá vỡ sự ổn định trong hệ thống khác. Thuyết chức năng của Merton thách thức chúng ta chỉ ra tất cả các chức năng cũng như phản chức năng cho các hệ thống khác nhau.

Đối với các chức năng của hệ thống, ông còn phân biệt các chức năng biểu hiện (manifest function) và các chức năng tiềm ẩn (latent function). Sự phân biệt cũng tương tự đối với các phản chức năng. So sánh giữa các chức năng và các phản chức năng, tức là so sánh những lực lượng tạo nên sự ổn định của hệ thống và những lực lượng phá vỡ sự ổn định của hệ thống đưa đến một khái niệm mới là hậu quả thuần túy đối với hệ thống (net aggregate of consequences). Nếu như các chức năng mạnh hơn các phản chức năng, điều đó có nghĩa là hành vi, ý tưởng hoặc đối tượng đang được nghiên cứu sẽ tồn tại dai dẳng. Bằng cách đưa ra những khái niệm mới này, thuyết chức năng của Merton không xem xung đột như là những yếu tố bên ngoài hệ thống. Điều quan trọng cần phải nhớ là ý nghĩa xã hội học đặc biệt của các chức năng và các phản chức năng. Kết luận rằng một cái gì đó có chức năng (dẫn đến sự ổn định của hệ thống) không có nghĩa rằng điều đó là tốt theo ý nghĩa đạo đức. Chúng ta phải có những chuẩn mực đạo đức và sử dụng những bằng chứng của phân tích chức năng để thay đổi sự vật và phản ánh những phê phán đạo đức một cách têt nhất. Không thể thay đổi trật tự của sự vật đơn thuần chỉ bằng sự không tán thành về n ạt đạo đức. Những tri thức về sự hoạt động của hệ thống sẽ làm tăng thêm yếu tố phê phán đối với khả năng của chúng ta trong việc cải tạo thế giới.

Việc ứng dụng mô hình chức năng theo quan niệm của Merton nghiên cứu những hiện tượng xã hội như sự nghèo khổ, tội phạm, nạn mãi dâm v.v... sẽ chỉ ra quan hệ có tính chất cấu trúc của hiện tượng đang nghiên cứu với các hiện tượng xã hội khác, chỉ ra những nguyên nhân. Qua đó con người có thể tác động đến những yếu tố nào trong cấu trúc của hệ thống,

hướng của sự tác động, tạo nên sự biến đổi theo hướng mong muốn của con người. Bởi vì bất cứ một biến đổi xã hội nào, sự phát triển nào cũng đều dẫn tới sự biến đổi có tính chất cấu trúc

Hướng dẫn cách đọc kết quả xử lý số liệu điều tra xã hội học bằng chương trình SPSS/PC+

TÔN LƯƠNG CHÍNH

I. Sơ lược về Chương trình SPSS/PC+

Ngày nay, các cuộc điều tra Xã hội học mang tính định lượng thường được xử lý trên máy vi tính bằng chương trình (phần mềm) SPSS/PC+. SPSS là từ viết tắt tiếng Anh của Statistical Package For Social Science, nghĩa là Bộ Chương trình thống kê cho Khoa học xã hội.

Chương trình SPSS có khả năng tính được tần số, tần suất, trung vị, mốt, trung bình số học, phương sai, sai số tiêu chuẩn... của từng biến riêng lẻ của toàn mẫu cũng như của từng phân nhóm trong mẫu một cách đồng thời. Ví dụ cùng một lúc ta có thể tính được tuổi kết hôn trung bình của những người được phỏng vấn và tuổi kết hôn trung bình của nhóm Nam và Nữ, hoặc tuổi kết hôn của những người ở các độ tuổi khác nhau... SPSS cho phép ta tính được tương quan (mối liên hệ) của hai biến. Điểm ưu việt của chương trình này là ở chỗ: nó không chỉ cung cấp cho chúng ta bảng tương quan giữa hai biến (hoặc ba biến) mà còn cung cấp cho chúng ta hệ số tương quan, mà các chương trình khác như FOX, FOXBASE, LOTUS... không có được. Hệ số tương quan biểu thị mối liên hệ mạnh hay yếu của hai biến. Thông qua hệ số tương quan này chúng ta có thể biết được tác động của các biến khác nhau đến một biến nào đó (chẳng hạn xem tác động của biến học vấn và biến mức sống đến biến đánh giá chính sách đổi mới) thì biến nào mạnh hơn. Hệ số tương quan có đặc tính chung là chúng biến thiên trong khoảng từ 0 đến 1. Khi hệ số tương quan bằng 0 giữa hai biến không có mối liên hệ. Khi nó bằng 1 thì giữa hai biến có mối liên hệ hàm số (mối liên hệ rất chặt). nghĩa là ta có thể biểu thị mối liên hệ đó bằng số, chẳng hạn bằng hàm số tuyến tính có dạng:

$$y = ax + b$$

Trong đó x, y là các biến cần đo, còn a, b là các hằng số. Khi hệ số tương quan khác 0 thì ta nói rằng giữa hai biến có mối liên hệ tương quan. Hệ số tương quan càng lớn (càng gần li) mối liên hệ càng chặt. Nói cách khác ảnh hưởng của biến số độc lập đến biến số phụ thuộc càng mạnh. Chẳng hạn hệ số tương quan giữa biến học vấn và biến đánh giá chính sách đổi mới bằng 0.25 còn giữa biến mức sống và biến đánh giá chính sách đổi mới là 0.40 thì lúc đó ta có thể kết luận được rằng: học vấn và mức sống của người dân có ảnh hưởng đến cách đánh giá của họ về chính sách đổi mới, song mức sống của những người được hỏi có ảnh hưởng mạnh đến nhận thức của họ về chính sách đổi mới hơn học vấn của họ.

Ngoài ra đối với hai biến định lượng, chương trình SPSS còn tính được mối liên hệ hồi quy tuyến tính, nghĩa là đưa ra được sự phụ thuộc dạng hàm số của hai biến này. Một ưu điểm khác của chương trình SPSS là ở chỗ nó có thể tạo ra cho ta các biến mới từ hai hoặc ba biến trở lên đã có sẵn bằng các phép tính số học, lôgic.

Chương trình SPSS là chương trình sử dụng các công cụ thống kê - toán, bởi vậy các tính toán của nó được thực hiện trên các ngôn ngữ số. Chính vì lẽ đó các thông tin sơ cấp thu được từ các cuộc điều tra Xã hội học bằng phương pháp phỏng vấn tiêu chuẩn (có bản hỏi), ăng kết cần phải được chuyển qua ngôn ngữ số. Công việc chuyển các thông tin bằng lời sang thông tin bằng số được gọi là mã hóa (code). Các thông tin bằng số trong ngôn ngữ máy được gọi là các giá trị (Value) còn thông tin bằng lời ứng với các giá trị đó là tên của giá trị (Value Label). Sau khi mã hóa (chuyển ngôn ngữ lời sang ngôn ngữ số) các thông tin này sẽ được chương trình SPSS/PC+ xử lý và phân tích theo yêu cầu của người nghiên cứu đối với từng câu hỏi (từng biến) cụ thể nhằm đáp ứng mục tiêu đặt ra cho từng câu hỏi phù hợp với dạng thang đo của câu hỏi đó.

Thông thường, chương trình SPSS cho ta các dạng kết quả sau:

1- Bảng tần số, tần suất (Frequency) của từng biến riêng lẻ. Khi tính toán tần số (Frequency) chương trình SPSS còn có thể cho ta biết thêm một số thông số hỗ trợ như: Trung bình số học (Mean), Mốt (Mode): giá trị có tần số cao nhất, Trung vị (Median): giá trị phân chia mau thành hai phần bằng nhau theo biến đang tính, Phương sai, Độ lệch chuẩn... Ngoài ra khi tính tần số máy cũng có thể vẽ được đồ thị phân bố của các giá trị theo kiểu hình cột hoặc đa giác.

2- Bảng giá trị trung bình số học (Mean) của một biến nào đó (thường là các biến đo bằng thang đo định lượng, ví dụ: tiền lương, tuổi, số thóc thu được...) với một biến nào đó. Chẳng hạn với câu hỏi: "Trung bình mỗi năm gia đình ta thu được bao nhiêu thóc". Lúc đó bảng này cho ta không những số thóc trung bình của toàn mẫu, mà còn có thể của từng nhóm xã hội nhỏ trong mẫu theo những tiêu chí (biến) khác nhau mà ta cần tính cho nó như: mức sống, quy mô gia đình, nghề nghiệp của gia đình... Bên cạnh đó khi tính trung bình số học Chương trình còn cho ta thêm thông tin về độ lệch tiêu chuẩn (phương sai) theo từng phân nhóm và số người (case) của phân nhóm được đưa vào tính giá trị trung bình số học.

3- Bảng tương quan của hai biến số. Bảng này thường có dạng bảng hai chiều với k hàng và l cột. Số lượng hàng và cột phụ thuộc vào số lượng các phương án trả lời của các biến mà ta cần tính tương quan. Ví dụ: Biến A có 3 phương án trả lời, còn biến B có 5 phương án trả lời thì số $k=3$ còn số $l=5$. Chương trình SPSS chỉ tính tương quan cho hai biến là hai câu hỏi tuyến (mỗi người chỉ được chọn 1 phương án trả lời trong số phương án vạch ra) chứ không tính được cho các câu hỏi hội. Trong bảng ngoài số lượng tuyệt đối (tần số), còn có số lượng tương đối (phần trăm). Số phần trăm có thể tính được theo hàng (Row Percent), theo cột (Column Percent) hoặc cả hai tùy theo yêu cầu và ý nghĩa của cách tính.

Ở cuối bảng có đưa ra hệ số tương quan của hai biến. Có nhiều loại hệ số tương quan khác nhau như: Phi, Cramer, Persson... Tùy theo từng dạng thang đo của các biến mà chúng ta yêu cầu máy tính hệ số thuộc dạng nào cho phù hợp. Thông thường hệ số tương quan Cramer hay được dùng nhất, bởi vì hệ số này dùng được cho tất cả các dạng thang đo, từ thang định tính đến thang đo định lượng.

II- Hướng dẫn sử dụng các kết quả.*1- Bảng Tần số (Frequency)*

Bảng tần số có dạng sau:

A1 Thôn điều tra

Value Label	Value	Frequency	Percent	Valid Percent	Cum Percent
Xuân Viên	1	106	52.0	52.5	52.5
Mẽ Sơn	2	61	30.2	30.2	82.7
Đông Sơn	3	17	8.4	8.4	91.1
Xuân Cầm	4	18	9.9	8.9	100.0
		1	5	MISSING	
	TOTAL	203	100.0	100.0	
Mean	Std Err	Median
Mode	Std Dev	Variance
Valid Cases	202	Missing Cases	1		

Giải thích:

Chú ý: Chương trình này được viết bằng tiếng Anh chỉ nên các phần giải thích bằng tiếng Việt như. Nội dung câu hỏi, các phương án trả lời đều không có dấu, mặt khác trong tiếng Anh quy ước chữ số có khác với tiếng Việt: dấu phẩy thập phân được thay bằng dấu chấm cho nên cần lưu ý khi đọc kết quả. Ví dụ: 34.6 cần hiểu là 34,6.

- **Value Label:** Tên của giá trị, nghĩa là tên của phương án trả lời trong câu hỏi mà ta đang tính. Ví dụ: Tên của thôn trong câu hỏi: "Hiện nay gia đình ta sống ở thôn nào"

- **Value:** Chữ số, mà ta gán mã cho tên phương án đó (mã hóa) để đưa vào máy tính. Frequency. Tần số, nghĩa là số người trả lời theo các phương án khác nhau của câu hỏi này. Ví dụ: Số người ở thôn Xuân Viên là 106 người.

- **Percent (Phần trăm):** Tỷ lệ phần trăm số người trả lời của mỗi phương án (Value) trên tổng số người được điều tra, kể cả những người không trả lời câu hỏi này. Ví dụ trong số 203 người được điều tra có 106 người ở thôn Xuân Viên, chiếm 52,0%

- **Valid Percent (Tỷ lệ phần trăm thực):** Tỷ lệ phần trăm số người trả lời của mỗi phương án (value) trên tổng số người trả lời câu hỏi này. Ví dụ, ở bảng trên do có 1 người không trả lời câu hỏi: "Sống ở thôn nào nên tỷ lệ phần trăm thực (Valid Percent) của số người ở thôn Xuân Viên (Value=1) là 52,3%

- **Cum Percent (Tần suất tích lũy hay phần trăm tích lũy):** Tổng phần trăm của các phương án trả lời đến phương án tương ứng. :

- **Missing:** Số đơn vị thiếu vắng không quan sát được (không trả lời câu hỏi này) .

- **Total:** Tổng số theo các mục (cột) . Ví dụ : Total theo cột Frequency là tổng số người được điều tra, bằng 203.

- **Mean:** Trung bình số học của biến đang tính (thường chỉ được tính cho các biến đo bằng thang đo định lượng như: Thu nhập, tuổi đời, số con...).

- **Std Err** : Sai số tiêu chuẩn.

- **Median**: Trung vị, điểm phân chia tập hợp mẫu (số người được điều tra) thành hai phần bằng nhau, nghĩa là có 50% số người được điều tra nằm ở dưới trung vị, còn 50% người còn lại có giá trị của biến (câu hỏi) đang tính lớn hơn trung vị.

- **Mode**: Giá trị của biến (phương án trả lời của câu hỏi) có tần số cao nhất. Mode cũng như Median là một dạng trung bình của các biến định tính.

- **Std Dev**: Độ lệch tiêu chuẩn (có tài liệu thống kê gọi là độ lệch quân phương), nói lên sự phân tán các giá trị của biến được tính so với giá trị trung bình. Ví dụ: 68% người trả lời có giá trị nằm trong khoảng từ trung bình số học (Mean) trừ đi độ lệch tiêu chuẩn (Std Dev) đến trung bình số học cộng với độ lệch tiêu chuẩn.

- **Variance**: Phương sai biểu thị độ phân tán của các giá trị quanh trung bình số học. Phương sai bằng bình phương của độ lệch tiêu chuẩn. (Để hiểu rõ hơn ý nghĩa của các con số này xin xem thêm tài liệu thống kê - toán).

- **Valid Cases** : Số người trả lời câu hỏi này.

- **Missing Cases** : Số người không trả lời câu hỏi này.

Tổng của Valid Cases và Missing Cases bằng số người được điều tra.

2- Bảng tương quan (Crosstabulation)

Biểu thị mối liên hệ của hai biến (hai câu hỏi hay hai dấu hiệu) với nhau. Thông thường là một biến độc lập và một biến phụ thuộc. Bảng tương quan có dạng sau:

Crosstabulation A5 Nội dung (tên) của biến A5 (câu hỏi A5)

by B1 Nội dung (tên) của biến B1 (câu hỏi B1)

Count Nội dung (tên) của các phương án trả lời của biến B1

Bl → Col Pet	1	2	3	Row Total
A5	1	3	2	5
Nội dung	27.3	10.0		14.7
(tên) của	2	5	1	6
các phương án	45.5	5.0		17.6
trả lời	3	3	17	23
	27.3	85.0	100.0	67.6
Cau A5 Column	11	20	3	34
Total	34.4	58.8	8.8	100.0

Statistic Value Significance

Cramer"s V 43876

Number of Missing Observations = 168
 Number of Missing Observations = 168

Giải thích:

Mở đầu bảng có chữ **Crosstabulation** (tương quan), tên biến thứ nhất (thường được biểu thị bằng chữ cái với chữ số viết tắt tên biến, ví dụ chữ **A1**) tiếp đó là giải thích đầy đủ tên biến (nội dung câu hỏi thứ nhất).

Dòng thứ hai có chữ **By**, tiếp đó là viết tắt tên của biến thứ hai (ví dụ B1) và giải thích đầy đủ của biến thứ hai (nội dung câu hỏi thứ hai). Chú ý: Giải thích đầy đủ tên biến không được dài quá 48 ký tự cho nên đối với những câu hỏi dài chỉ có thể viết được các chữ đầu; Giải thích các phương án trả lời không được dài quá 16 ký tự (kể cả khoảng cách giữa hai từ), dài quá số quy định ở trên máy sẽ tự động cắt. Toàn bộ hai dòng đầu này có ý nghĩa là bảng tương quan của biến số nào (câu hỏi nào) với biến (câu hỏi) nào.

Tiếp đến là Bảng số liệu tương quan của hai biến số. Các phương án trả lời của biến thứ hai được bố trí theo hàng ngang của dòng đầu (có mũi tên chỉ vào, ví dụ B1) còn các phương án trả lời của biến thứ nhất được bố trí trong cột đầu. Trong phần các phương án trả lời có nội dung các phương án trả lời và chữ số mã của nó. Trên cùng của cột đầu có ghi chữ **Count**, nghĩa là cách tính phần trăm như thế nào. Nếu dưới chữ Count có chữ:

- **Row. Pct**: Tính phần trăm theo hàng, nghĩa là bằng số ở ô đó chia cho tổng số các số trong hàng cùng với ô đó nhân với 100.

- **Col. Pct**: Tính phần trăm theo cột.

Hoặc có cả hai chữ: **Row. Pct** và **Col. Pct** tức là tính phần trăm cho cả hàng lẫn cột

Trong mỗi ô của bảng tương quan có hai hoặc ba chữ số. Chữ số ở trên (số nguyên) là tần số, chữ số thứ hai (số thập phân) là tỷ lệ phần trăm, chữ số thứ ba (nếu có) cũng là chữ số thập phân, là tỷ lệ phần trăm.

Cột cuối cùng bên tay phải của bảng có chữ **Row Total**, nghĩa là cách trả lời các phương án của tất cả mẫu đối với biến A1.

Dòng cuối cùng của cột đầu có chữ **Column Total**, cách trả lời các phương án đối với câu B1 của tất cả các mẫu (cả cuộc điều tra).

Tiếp theo dưới bảng có các dòng chữ sau:

Dòng thứ nhất:

- **Statistic**: tên hệ số tương quan.

- **Value**: giá trị của hệ số tương quan.

Dòng thứ hai: Dưới chữ Statistic là tên hệ số tương quan mà ta tính, chẳng hạn: Crammer, Phi, Student..., dưới chữ Value là giá trị của hệ số tương quan.

Dòng cuối cùng của bảng này là chữ:

Number of Missing Observation : Nghĩa là tổng số người không có thông tin một trong hai biến. Ví dụ: câu A1 có 3 người không trả lời, còn câu B1 có 5 người không trả lại thì Number of Missing Observation bằng 8.

3- Bảng tính giá trị trung bình số học

Bảng trung hình số học của một biến nào đó tính cho các nhóm nhỏ trong mẫu (theo cách phân chia nào đó toàn mẫu thành nhóm nhỏ) thường có dạng sau đây:

Summaries of A2

Tên biến cần tính trung bình

By levels of A1

Tên biến lấy làm chuẩn để phân chia thành nhóm nhỏ.

Variable	Value Label	Mean	Std Dev	Cases
For Entire	Population	52.6190	55.0671	126
Al	1 (Tên các nhóm)	58.1875	22.9662	32
Al	3 nhỏ của tiêu	49.7727	14.3093	22
Al	4 chuẩn	40.3448	18.2860	29
Al	5 Phân chia	156.8000	263.4838	5
Al	6 Al)	48.5238	15.6257	21
Al	7	41.1765	15.9501	17

Total Cases = 202

Missing Case8 = 76 OR 37.6 PCT.

Ở đây cột Variable cột nêu lên biến (câu hỏi, dấu hiệu) làm tiêu chuẩn phân chia mẫu thành các nhóm nhỏ. Cột Value Labels là cột tên các nhóm nhỏ, còn cột Mean là giá trị trung bình của từng nhóm nhỏ, cột Std Dev là độ lệch tiêu chuẩn của nhóm và cuối cùng cột Cases là số người được tính trong mỗi nhóm.

Dưới dòng chữ tiếng Anh đó có dòng chữ For Entire Population nghĩa là trung bình cho toàn mẫu.

Total Cases : Số người được điều tra.

Missing cases : Tổng số người không trả lời của hai câu là bao nhiêu và chiếm bao nhiêu phần trăm.