

# CHƯƠNG 2: PHƯƠNG PHÁP

## 2.1. NGUỒN SỐ LIỆU

Báo cáo chuyên khảo này sử dụng số liệu điều tra mẫu 15% của Tổng điều tra dân số và nhà ở Việt Nam năm 2009, mẫu 3% của cuộc Tổng điều tra dân số và nhà ở năm 1999 và mẫu 5% của cuộc Tổng điều tra dân số và nhà ở năm 1989. Báo cáo cũng sử dụng các thông tin từ các ấn phẩm của Tổng điều tra dân số và các nguồn số liệu khác bao gồm số liệu từ các vòng điều tra gần đây nhất của Điều tra mức sống hộ gia đình trong các năm 2006 và 2008[5]. Một số số liệu chỉ tiêu và ước lượng ở cấp tỉnh về nhân khẩu học (TFR, SRB) đã được tính toán từ ĐTDS2009 cũng được sử dụng trong báo cáo chuyên khảo này.

Để trình bày bằng hình ảnh kết quả một số phân tích tương quan, số liệu về hệ thống thông tin địa lý (GIS) của Việt Nam cũng được sử dụng, dựa trên thông tin định dạng cơ bản của 63 đơn vị hành chính của Việt Nam.

## 2.2. ĐỊNH NGHĨA CÁC CHỈ TIÊU/BIẾN SỐ

Các chỉ tiêu được xem xét trong báo cáo chuyên khảo này bao gồm một số chỉ tiêu cơ bản về nhân khẩu học và kinh tế-xã hội, các chỉ tiêu liên quan đến giáo dục và đào tạo. Tổng cục Thống kê và Ban Chỉ đạo Tổng điều tra dân số và nhà ở Trung ương đã đưa ra những khái niệm/định nghĩa chi tiết về các chỉ tiêu này [3, 6]. Một số khái niệm/định nghĩa có thể tóm tắt như sau:

- (1) Tình trạng biết đọc biết viết: Một người được coi là biết đọc biết viết nếu người đó có thể đọc, viết và hiểu đầy đủ những câu đơn giản bằng chữ quốc ngữ, chữ dân tộc hoặc chữ nước ngoài. Tương tự, một người được coi là không biết đọc biết viết nếu không thể làm được những việc trên.
- (2) Tỷ lệ dân số từ 15 tuổi trở lên biết đọc, biết viết là tỷ lệ phần trăm những người từ 15 tuổi trở lên biết đọc biết viết trên tổng số người từ 15 tuổi trở lên.
- (3) Trình độ học vấn chia theo các cấp học
  - Chưa bao giờ đến trường: là người chưa từng đi học ở các trường, lớp nào thuộc Hệ thống giáo dục quốc dân.
  - Chưa tốt nghiệp tiểu học: là người đã từng đi học nhưng chưa tốt nghiệp tiểu học, kể cả người học lớp cuối cùng của bậc tiểu học nhưng chưa đỗ tốt nghiệp.
  - Tốt nghiệp tiểu học: là người đã từng đi học và tốt nghiệp tiểu học, kể cả người đã từng học trung học cơ sở nhưng chưa đỗ tốt nghiệp bậc học này.
  - Tốt nghiệp trung học cơ sở: là người đã từng đi học và tốt nghiệp trung học cơ sở, kể cả người đã từng học trung học phổ thông nhưng chưa đỗ tốt nghiệp bậc học này.

- Tốt nghiệp trung học phổ thông trở lên: là người đã từng đi học và tốt nghiệp trung học phổ thông, hoặc đã hay chưa tốt nghiệp các bậc học cao hơn trung học phổ thông như cao đẳng, đại học, thạc sĩ, tiến sĩ.
- (4) Tỷ lệ dân số 5-18 tuổi bỏ học: là tỷ lệ phần trăm những trẻ trong độ tuổi 5-18 hiện không đi học trên tổng số trẻ từ 5-18 tuổi.<sup>1</sup>
- (5) Tỷ lệ nhập học đúng tuổi các cấp:
- Tỷ lệ nhập học đúng tuổi cấp tiểu học là tỷ lệ phần trăm số học sinh trong độ tuổi từ 6 đến 11 tuổi đang tham gia vào giáo dục tiểu học trong tổng dân số trong độ tuổi học tiểu học (từ lớp 1 đến lớp 5 và tuổi của học sinh vào học lớp 1 là 6 tuổi).
  - Tỷ lệ nhập học đúng tuổi cấp trung học cơ sở là tỷ lệ phần trăm số học sinh trong độ tuổi từ 12 đến 15 tuổi đang tham gia vào giáo dục trung học cơ sở trong tổng dân số trong độ tuổi học trung học cơ sở (từ lớp 6 đến lớp 9 và học sinh vào học lớp 6 phải hoàn thành chương trình tiểu học và có tuổi là 11 tuổi).
  - Tỷ lệ nhập học đúng tuổi cấp trung học phổ thông là tỷ lệ phần trăm số học sinh trong độ tuổi từ 16 đến 18 tuổi đang tham gia vào giáo dục trung học phổ thông trong tổng dân số trong độ tuổi học trung học phổ thông (từ lớp 10 đến lớp 12 và học sinh vào học lớp 10 phải có bằng trung học cơ sở và có tuổi là 15 tuổi).
- (6) Trình độ chuyên môn kỹ thuật cao nhất đã đạt được:
- Một người được coi là có trình độ Sơ cấp nếu trình độ chuyên môn kỹ thuật cao nhất đã được đào tạo và được cấp chứng chỉ là sơ cấp hoặc dạy nghề dưới 3 tháng.
  - Một người được coi là có trình độ Trung cấp nếu trình độ chuyên môn kỹ thuật cao nhất đã được đào tạo và được cấp bằng là trung cấp chuyên nghiệp hoặc trung cấp nghề.
  - Một người được coi là có trình độ Cao đẳng nếu trình độ chuyên môn kỹ thuật cao nhất đã được đào tạo và được cấp bằng là cao đẳng hoặc cao đẳng nghề.
  - Một người được coi là có trình độ Đại học trở lên nếu người đó đã được đào tạo và được cấp bằng đại học, thạc sĩ hoặc tiến sĩ.

Bên cạnh phân loại trình độ chuyên môn kỹ thuật trên, hệ thống Phân loại chuẩn quốc tế về giáo dục (ISCED) do UNESCO xây dựng cũng được sử dụng để so sánh các chỉ số về trình độ giáo dục giữa Việt Nam và các nước khác. Thông tin thêm về hệ thống phân loại này có trong Phụ lục 3 và người đọc có thể tham khảo thêm thông tin lý giải đầy đủ về hệ thống phân loại này ở tài liệu khác [7]. Phiên bản rút gọn của hệ thống phân loại này được nhóm biên soạn chuyên khảo sử dụng, trong đó trình độ học vấn bậc thấp là thuật ngữ chỉ các trình độ tiểu học và dưới tiểu học (ISCED 1) và trung học cơ sở (ISCED 2); trình độ học vấn bậc trung là thuật ngữ chỉ trình độ trung học phổ thông (ISCED 3) và trình độ trên trung học phổ thông nhưng dưới đại học (ISCED 4); trình độ học vấn bậc cao là thuật ngữ chỉ trình độ đại học và trên đại học (ISCED 5 và 6). Cần lưu ý rằng các chỉ số này được tính toán cho nhóm dân số từ 25 tuổi trở lên.

<sup>1</sup> Lưu ý rằng số trẻ em trong độ tuổi 5-18 hiện không đi học có thể bao gồm cả những trẻ trong độ tuổi này đã hoàn thành sớm toàn bộ chương trình giáo dục phổ thông.

(7) Một số chỉ tiêu liên quan đến mức sinh:

- Tổng tỷ suất sinh (TFR) trong vòng mười hai tháng qua là số con sinh sống bình quân của một phụ nữ trong cả cuộc đời, nếu người phụ nữ đó trong suốt thời kỳ sinh đẻ tuân theo tỷ suất sinh đặc trưng theo tuổi như quan sát được trong 12 tháng trước khi thực hiện cuộc Tổng điều tra dân số và nhà ở năm 2009. Báo cáo sử dụng tổng tỷ suất sinh của 63 tỉnh/thành phố của Việt Nam đã được công bố trong báo cáo “Tổng điều tra dân số và nhà ở Việt Nam năm 2009 – Các kết quả chủ yếu” [3].
- Tỷ lệ phụ nữ sinh con thứ ba trở lên là tỷ lệ phần trăm phụ nữ sinh con thứ ba trở lên trong vòng 12 tháng trước thời điểm Tổng điều tra trên tổng số phụ nữ có tuổi từ 15 đến 49 đã sinh con trong thời kỳ tham chiếu.
- Tỷ số giới tính được biểu thị bằng số nam trên 100 nữ của dân số. Tỷ số giới tính của trẻ em dưới 5 tuổi sử dụng kết quả tính toán của chuyên khảo tỷ số giới tính khi sinh.

(8) Một số chỉ tiêu về di cư:

- Người không di cư là những người đã sinh sống ở cùng một xã trong 5 năm trước thời điểm Tổng điều tra dân số.
- Người di cư trong tỉnh là những người đã di chuyển từ một địa phương khác, nhưng cùng trong một tỉnh với nơi hiện đang sinh sống, tới nơi mà cuộc phỏng vấn điều tra được thực hiện.
- Người di cư giữa các tỉnh là những người đã từng sống tại một tỉnh khác (hoặc ở nước ngoài) 5 năm trước thời điểm Tổng điều tra dân số.

(9) Một số chỉ tiêu về điều kiện kinh tế bao gồm: thu nhập bình quân đầu người hàng tháng và tỷ lệ hộ nghèo. Báo cáo sử dụng thu nhập bình quân đầu người hàng tháng được công bố trong báo cáo của điều tra mức sống hộ gia đình năm 2008[5], số liệu tỷ lệ hộ nghèo được tính cho từng tỉnh/thành phố từ số liệu của Điều tra mức sống hộ gia đình năm 2006 [8]. Chuẩn nghèo của chính phủ thời kỳ 2006-2010 là 200 nghìn đồng/người/tháng đối với nông thôn và 260 nghìn đồng/ người/tháng đối với khu vực thành thị.

(10) Tình trạng giàu nghèo của hộ gia đình: Tổng điều tra dân số và nhà ở Việt Nam năm 2009 không thu thập thông tin về thu nhập của hộ gia đình và từng cá nhân. Tuy nhiên một số chỉ tiêu về điều kiện sống và sinh hoạt của hộ gia đình đã được thu thập như: thông tin về nhà ở (kết cấu nhà, diện tích ở, v.v), điều kiện sinh hoạt (nguồn nước, nguồn điện, hố xí, v.v.) và trang thiết bị trong hộ gia đình (tivi, tủ lạnh, xe máy, v.v.). Từ những thông tin này, phương pháp phân tích nhân tố (factor analysis) được sử dụng để tính ra một chỉ số gián tiếp đo lường điều kiện kinh tế xã hội của hộ gia đình. Chỉ số gián tiếp được chuyển thành phân nhóm ngũ phân vị (quintile) của điều kiện kinh tế-xã hội hộ gia đình thông qua việc sử dụng các điểm cắt các mức giá trị thứ 20%, 40%, 60% và 80% của toàn bộ mẫu nghiên cứu. Tương ứng, năm nhóm mức độ kinh tế hộ gia đình được phân ra là: nghèo nhất, nghèo, trung bình, giàu, và giàu nhất. Lưu ý rằng mặc dù phương pháp phân tích nhân tố áp dụng thông tin về nhà ở, điều kiện sinh hoạt và trang thiết bị hộ gia đình khá phổ biến trong các cuộc điều tra ở các nước đang phát triển, chỉ số kinh tế-xã hội do phương pháp này đưa ra chỉ là một chỉ số gián tiếp, không tương ứng với chuẩn nghèo hiện tại của chính phủ Việt Nam.

## 2.3. PHƯƠNG PHÁP PHÂN TÍCH SỐ LIỆU

Các số liệu sử dụng trong phân tích là các số liệu đã được suy rộng (gia quyền).

Kỹ thuật lập bảng chéo cũng được sử dụng để tìm hiểu sự khác biệt về tất cả các chỉ số giáo dục và đào tạo cơ bản theo các yếu tố nhân khẩu học và kinh tế-xã hội. Các phân tích tương quan và lập bản đồ được thực hiện trên số liệu tổng hợp cho 63 tỉnh/thành phố.

## 2.4. PHẦN MỀM PHÂN TÍCH SỐ LIỆU

Phần mềm SPSS phiên bản 18.0 được sử dụng để tổng hợp số liệu, phân tích đơn biến và đa biến. Phần mềm thông tin địa lý ArcGIS phiên bản 9.3 được sử dụng để gán số liệu cấp tỉnh/thành phố cũng như thể hiện sự tương quan trên bản đồ Việt Nam.

Phần mềm Spectrum 3.4 được sử dụng để đưa ra một số dự báo bằng phương pháp dự báo thành phần (component projection) cho một số chỉ tiêu giáo dục (như số lượng học sinh, số lượng lớp học và giáo viên cần có). Các giả thuyết đầu vào để tiến hành dự báo như sau:

- (1) Dân số theo giới tính và nhóm 5 độ tuổi: Sử dụng kết quả toàn bộ Tổng điều tra dân số và nhà ở 0 giờ ngày 1 tháng 4 năm 2009.
- (2) Mức sinh: Tổng tỷ suất sinh (TFR)
  - 2009: 2,03 con/phụ nữ;
  - 2024: 1,80 con/phụ nữ và giữ nguyên mức này cho đến năm 2059.
- (3) Mức chết: Tuổi thọ trung bình (e0)
  - Năm 2009: 70,2 tuổi đối với nam và 75,6 tuổi đối với nữ;
  - Năm 2030: 75,4 tuổi đối với nam và 80,0 tuổi đối với nữ và giữ nguyên đến cuối kỳ dự báo;
- (4) Di cư: Giả thuyết di cư bằng 0.

## 2.5. HẠN CHẾ

Báo cáo chuyên khảo này phân tích số liệu từ nhiều nguồn khác nhau. Các phân tích và đánh giá các mối quan hệ sử dụng các chỉ số tổng hợp đã được công bố. Một hạn chế nữa là việc tách, nhập các đơn vị hành chính trong suốt hơn 20 năm qua gây ra một số khó khăn trong việc phân loại chính xác các đơn vị phân tích, nên có thể có một số sai lệch nhỏ trong các chỉ tiêu tính toán.

Mặc dù mẫu ĐĐTDS năm 2009 đã lớn hơn khá nhiều so với Tổng điều tra năm 1989 và Tổng điều tra năm 1999, tuy nhiên khi phân tích theo nhóm dân tộc ít người thì cỡ mẫu rất nhỏ, không có tính đại diện, nhất là khi phân tích trên số liệu chia theo tuổi và giới tính ở các vùng kinh tế - xã hội và tỉnh/thành phố.