

BỘ NÔNG NGHIỆP VÀ PHÁT TRIỂN NÔNG THÔN
TRƯỜNG CAO ĐẲNG CƠ GIỚI NINH BÌNH

GIÁO TRÌNH

MÔN HỌC: THỐNG KÊ XÃ HỘI

NGHỀ: CÔNG TÁC XÃ HỘI

TRÌNH ĐỘ: CAO ĐẲNG NGHỀ

*Ban hành kèm theo Quyết định số: /QĐ-TCDCGNB ngày.....tháng....năm
2017*

của Hiệu trưởng Trường cao đẳng Cơ giới Ninh Bình

Ninh Bình, năm 2017

TaiLieu.vn

TUYÊN BỐ BẢN QUYỀN

Tài liệu này thuộc loại sách giáo trình nên các nguồn thông tin có thể được phép dùng nguyên bản hoặc trích dùng cho các mục đích về đào tạo và tham khảo.

Mọi mục đích khác mang tính lệch lạc hoặc sử dụng với mục đích kinh doanh thiếu lành mạnh sẽ bị nghiêm cấm.

LỜI NÓI ĐẦU

Ngày nay, thống kê được coi là một trong những công cụ quản lý vĩ mô quan trọng, có vai trò cung cấp các thông tin thống kê trung thực, khách quan, chính xác, đầy đủ, kịp thời phục vụ các cơ quan nhà nước trong việc đánh giá, dự báo tình hình, hoạch định chiến lược, chính sách, xây dựng kế hoạch, chiến lược và chính sách phát triển kinh tế - xã hội ngắn hạn và dài hạn, trong phạm vi một lĩnh vực hay toàn bộ nền kinh tế, trong phạm vi một xã hay quốc gia. Đồng thời, các con số thống kê cũng là những cơ sở quan trọng nhất để kiểm điểm, đánh giá tình hình thực hiện các kế hoạch, chiến lược và các chính sách đó.

Nhằm kịp thời đáp ứng yêu cầu đào tạo nghề Công tác xã hội. Chúng tôi đã tiến hành biên soạn cuốn sách này làm tài liệu học tập cho các lớp đào tạo nghề Công tác xã hội.

Giáo trình được biên soạn theo Chương trình môn học Thống kê xã hội đã được phê duyệt. Cuốn sách gồm 5 chương:

Chương 1: Đối tượng nghiên cứu của thống kê xã hội

Chương 2: Quá trình nghiên cứu của thống kê

Chương 3: Các mức độ của hiện tượng kinh tế - xã hội

Chương 4: Nghiên cứu sự biến động của hiện tượng kinh tế xã hội

Chương 5: Phương pháp chọn mẫu

Đây là giáo trình biên soạn lần đầu tiên vì vậy không tránh khỏi những thiếu sót, tôi mong nhận được những ý kiến đóng góp quý báu của các nhà khoa học, nhà quản lý và đông đảo bạn đọc.

Trân trọng cảm ơn!

Ninh Bình, năm 2017

Tham gia biên soạn

1. Chủ biên: Vũ Ánh Dương

2. Lê Hùng Cường

TaiLieu.vn

MỤC LỤC

GIÁO TRÌNH MÔN HỌC

Tên môn học: Thống kê xã hội

Mã môn học: MH09

Vị trí, tính chất, ý nghĩa và vai trò của môn học:

- Vị trí môn học: Thống kê xã hội là môn lý thuyết kỹ thuật cơ sở quan trọng của chương trình đào tạo nghề công tác xã hội, liên quan đến nội dung công việc của nhân viên xã hội.

- Tính chất môn học: là môn học kỹ thuật cơ sở bắt buộc.

Mục tiêu môn học:

- Kiến thức: Trình bày được các khái niệm cơ bản của thống kê xã hội, quá trình nghiên cứu thống kê, các phương pháp thống kê, dự báo thống kê xã hội

- Kỹ năng: Vận dụng được các phương pháp thống kê, dự báo thống kê trong nghiên cứu các vấn đề xã hội để phân tích, tổng hợp số liệu, viết báo cáo...

- Thái độ: Cẩn thận, tỉ mỉ, chính xác, khách quan, linh hoạt, sáng tạo trong thống kê xã hội.

Nội dung môn học:

CHƯƠNG 1

ĐỐI TƯỢNG NGHIÊN CỨU CỦA THỐNG KÊ XÃ HỘI

Mã chương: MH10_CH01

Giới thiệu: Giới thiệu tổng quan về môn học và các khái niệm cơ bản, cũng như đối tượng nghiên cứu của thống kê xã hội. Giúp người học có cái nhìn tổng quát và những yêu cầu của môn học.

Mục tiêu:

- Kiến thức: Trình bày được nguồn gốc thống kê, các khái niệm cơ bản về thống kê xã hội, mục đích, ý nghĩa của thống kê xã hội.

- Kỹ năng: Áp dụng được các khái niệm, các loại thang đo, thu thập thông tin trong thống kê xã hội.

- Thái độ: Cẩn thận, tỉ mỉ, khách quan, chính xác trong quá trình thống kê.

Nội dung chính:

1. Nguồn gốc môn học

Nếu thống kê được hiểu theo nghĩa thông thường thì ngay từ thời cổ đại con người đã đã chú ý đến việc này thông qua việc ghi chép đơn giản.

Cuối thế kỷ XVII, lực lượng sản xuất phát triển mạnh mẽ làm cho phương thức sản xuất của chủ nghĩa tư bản ra đời. Kinh tế hàng hóa phát triển dẫn đến các ngành sản xuất riêng biệt tăng thêm, phân công lao động xã hội ngày càng phát triển. Tính chất xã hội của sản xuất ngày càng cao, thị trường được mở rộng không chỉ trong một nước mà toàn thế giới. Để phục vụ cho mục đích kinh tế, chính trị và quân sự nhà nước tư bản và các chủ tư bản cần rất nhiều thông tin thường xuyên về thị trường, giá cả, sản xuất, nguyên liệu, dân số,... Do đó, công tác thống kê phát triển nhanh chóng. Chúng ta có thể đưa ra 3 nhóm tác giả được gọi là những người khai sáng cho ngành khoa học thống kê:

- Những người đầu tiên đưa ngành khoa học thống kê đi vào thực tiễn, đại diện cho những tác giả này là nhà kinh tế học người Đức H.Conhring (1606 - 1681), năm 1660 ông đã giảng dạy tại trường đại học Halmsted về phương pháp nghiên cứu hiện tượng xã hội dựa vào số liệu điều tra cụ thể.

- Với những thành quả của người đi trước, bổ sung hoàn chỉnh thành môn học chính thống, đại diện là William Petty, một nhà kinh tế học của người Anh, là tác giả cuốn “Sổ học chính trị” xuất bản năm 1682, một số tác phẩm có tính chất phân tích thống kê đầu tiên ra đời.

- Thống kê được gọi với nhiều tên khác nhau thời bấy giờ, sau đó năm 1759 một giáo sư người Đức, Achenwall (1719-1772) lần đầu tiên dùng danh từ “Statistics” (một thuật ngữ gốc La tinh “Status”, có nghĩa là Nhà nước hoặc trạng thái của hiện tượng) - sau này người ta dịch ra là “Thống kê”.

Kể từ đó, thống kê có sự phát triển rất mạnh mẽ và ngày càng hoàn thiện, gắn liền với nhiều nhà toán học - thống kê học nổi tiếng như: M.V.Lomonoxop (nga, 1711-1765), Laplace (Pháp, 1749-1827), I.Fisher, W.M.Pearsons,...

2. Thống kê là gì?

2.1. Định nghĩa

Thống kê là một hệ thống các phương pháp bao gồm thu thập, tổng hợp, trình bày số liệu, tính toán các đặc trưng của đối tượng nghiên cứu nhằm phục vụ cho quá trình phân tích, dự đoán và ra quyết định.

2.2. Chức năng của thống kê

Thống kê thường được phân thành 2 lĩnh vực:

- Thống kê mô tả (Descriptive statistics): là các phương pháp có liên quan đến việc thu thập số liệu, tóm tắt, trình bày, tính toán và mô tả các đặc trưng khác nhau để phản ánh một cách tổng quát đối tượng nghiên cứu.

- Thống kê suy luận (Inferential statistics): là bao gồm các phương pháp ước lượng các đặc trưng của tổng thể, phân tích mối liên hệ giữa các hiện tượng nghiên cứu, dự đoán hoặc ra quyết định trên cơ sở thông tin thu thập từ kết quả quan sát mẫu.

2.3. Phương pháp thống kê

- Thu thập và xử lý số liệu:

Số liệu thu thập thường rất nhiều và hỗn độn, các dữ liệu đó chưa đáp ứng cho quá trình nghiên cứu. Để có hình ảnh tổng quát về tổng thể nghiên cứu, số liệu thu thập phải được xử lý tổng hợp, trình bày, tính toán các số đo; kết quả có được sẽ giúp khái quát được đặc trưng của tổng thể.

- Nghiên cứu các hiện tượng trong hoàn cảnh không chắc chắn:

Trong thực tế, có nhiều hiện tượng mà thông tin liên quan đến đối tượng nghiên cứu không đầy đủ mặc dù người nghiên cứu đã có sự cố gắng. Ví dụ như nghiên cứu về nhu cầu của thị trường về một sản phẩm ở mức độ nào, tình trạng của nền kinh tế ra sao, để nắm được các thông tin này một cách rõ ràng quả là một điều không chắc chắn.

- Điều tra chọn mẫu:

Trong một số trường hợp để nghiên cứu toàn bộ tất cả các quan sát của tổng thể là một điều không hiệu quả, xét cả về tính kinh tế (chi phí, thời gian) và tính kịp thời, hoặc không thực hiện được. Chính điều này đã đặt ra cho thống kê xây dựng các phương pháp chỉ cần nghiên cứu một bộ phận của tổng thể mà có thể suy luận cho hiện tượng tổng quát mà vẫn đảm bảo độ tin cậy cho phép, đó là phương pháp điều tra chọn mẫu.

- Nghiên cứu mối liên hệ giữa các hiện tượng:

Giữa các hiện tượng nghiên cứu thường có mối liên hệ với nhau. Ví dụ như mối liên hệ giữa chi tiêu và thu nhập; mối liên hệ giữa lượng vốn vay và các yếu tố tác động đến lượng vốn vay như chi tiêu, thu nhập, trình độ học vấn; mối liên hệ giữa tốc độ phát triển với tốc độ phát triển của các ngành, lạm phát, tốc độ phát triển dân số,... Sự hiểu biết về mối liên hệ giữa các hiện tượng rất có ý nghĩa, phục vụ cho quá trình dự đoán.

- Dự đoán:

Dự đoán là một công việc cần thiết trong tất cả các lĩnh vực hoạt động.

Trong hoạt

động dự đoán người ta có thể chia ra thành nhiều loại:

(1). Dự đoán dựa vào định lượng và dựa vào định tính. Tuy nhiên, trong thống kê chúng ta chủ yếu xem xét về mặt định lượng với mục đích cung cấp cho những nhà quản lý có cái nhìn mang tính khoa học hơn và cụ thể hơn trước khi ra quyết định phù hợp.

(2). Dự đoán dựa vào nội suy và dựa vào ngoại suy.

- Dự đoán nội suy là chúng ta dựa vào bản chất của hiện tượng để suy luận, ví dụ như chúng ta xem xét một liên hệ giữa lượng sản phẩm sản xuất ra phụ thuộc các yếu tố đầu vào như vốn, lao động và trình độ khoa học kỹ thuật.

- Dự đoán dựa vào ngoại suy là chúng ta chỉ quan sát sự biến động của hiện tượng trong thực tế, tổng hợp lại thành qui luật và sử dụng qui luật này để suy luận, dự đoán sự phát triển của hiện tượng. Ví dụ như để đánh giá kết quả hoạt động của một công ty người ta xem xét kết quả hoạt động kinh doanh của họ qua nhiều năm.

Ngoài ra, người ta còn có thể phân chia dự báo thống kê ra thành nhiều loại khác.

3. Các khái niệm thường dùng trong thống kê

3.1. Tổng thể thống kê (Populations)

Tổng thể thống kê là tập hợp các đơn vị cá biệt về sự vật, hiện tượng trên cơ sở một đặc điểm chung nào đó cần được quan sát, phân tích mặt lượng của chúng. Các đơn vị, phần tử tạo nên hiện tượng được gọi là các đơn vị tổng thể.

Như vậy muốn xác định được một tổng thể thống kê, ta cần phải xác định được tất cả các đơn vị tổng thể của nó. Thực chất của việc xác định tổng thể thống kê là việc xác định các đơn vị tổng thể.

Trong nhiều trường hợp, các đơn vị của tổng thể được biểu hiện một cách rõ ràng, dễ xác định. Ta gọi nó là tổng thể bộ lộ. Ngược lại, một tổng thể mà các đơn vị của nó không được nhận biết một cách trực tiếp, ranh giới của tổng thể không rõ ràng được gọi là tổng thể tiềm ẩn.

Đối với tổng thể tiềm ẩn, việc tìm được đầy đủ, chính xác gặp nhiều khó khăn. Việc nhầm lẫn, bỏ sót các đơn vị trong tổng thể dễ xảy ra. Ví dụ

như tổng thể là những những mê nhạc cổ điển, tổng thể người mê tín dị đoan,...

3.2. Mẫu (Samples)

Mẫu là một bộ phận của tổng thể, đảm bảo được tính đại diện và được chọn ra để quan sát và dùng để suy diễn cho toàn bộ tổng thể. Như vậy, tất cả các phần tử của mẫu phải thuộc tổng thể, nhưng ngược lại các phần tử của tổng thể thì chưa chắc thuộc mẫu. Điều này tưởng chừng là đơn giản, tuy nhiên trong một số trường hợp việc xác định mẫu cũng có thể dẫn đến nhầm lẫn, đặc biệt là trong trường hợp tổng thể ta nghiên cứu là tổng thể tiềm ẩn.

Ngoài ra, chọn mẫu như thế nào để làm cơ sở suy diễn cho tổng thể, tức là mẫu phải mang tính đại diện cho tổng thể. Điều này thực sự không dễ dàng, ta chỉ cố gắng hạn chế tối đa sự sai biệt này mà thôi chứ không thể khắc phục được hoàn toàn.

3.3. Quan sát (Observations)

Là mỗi đơn vị của mẫu ; trong một số tài liệu còn được gọi là quan trắc.

3.4. Tiêu thức thống kê

Các đơn vị tổng thể thường có nhiều đặc điểm khác nhau, tuy nhiên trong thống kê người ta chỉ chọn một số đặc điểm để nghiên cứu, các đặc điểm này người ta gọi là tiêu thức thống kê. Như vậy, tiêu thức thống kê là khái niệm chỉ các đặc điểm của đơn vị tổng thể. Mỗi tiêu thức thống kê đều có các giá trị biểu hiện của nó, dựa vào sự biểu hiện của nó người ta chia ra làm hai loại:

a) Tiêu thức thuộc tính: là tiêu thức phản ánh loại hoặc tính chất của đơn vị. Ví dụ như ngành kinh doanh, nghề nghiệp,...

b) Tiêu thức số lượng: là đặc trưng của đơn vị tổng thể được thể hiện bằng con số. Ví dụ, năng suất của một loại cây trồng.

Tiêu thức số lượng được chia làm 2 loại:

- Loại rời rạc: là loại các giá trị có thể của nó là hữu hạn hay vô hạn và có thể đếm được.

- Loại liên tục: là loại mà giá trị của nó có thể nhận bất kỳ một trị số nào đó trong một khoảng nào đó.

3.5. Tham số tổng thể

Là giá trị quan sát được của tổng thể và dùng để mô tả đặc trưng của hiện tượng nghiên cứu. Trong xác suất thống kê toán chúng ta đã biết các tham số tổng thể như trung bình tổng thể (μ), tỷ lệ tổng thể (p), phương sai tổng thể

(²). Ngoài ra, trong quá trình nghiên cứu sâu môn thống kê chúng ta còn có thêm nhiều tham số tổng thể nữa như: tương quan tổng thể (ρ), hồi qui tuyến tính tổng thể,...

3.6. Tham số mẫu

Tham số mẫu là giá trị tính toán được của một mẫu và dùng để suy rộng cho tham số tổng thể. Đó là cách giải thích mang tính chất thông thường, còn đối với xác suất thống kê thì tham số mẫu là ước lượng điểm của tham số tổng thể, trong trường hợp chúng ta chưa biết tham số tổng thể chúng ta có thể sử dụng tham số mẫu để ước lượng tham số tổng thể. Chúng ta có thể liệt kê vài tham số mẫu như sau: trung bình mẫu ($\bar{x}$), tỷ lệ mẫu ($p^{\wedge}$), phương sai mẫu (S^2), hệ số tương quan mẫu (r),...

4. Các loại thang đo

Đứng trên quan điểm của nhà nghiên cứu, chúng ta cần xác định các phương pháp phân tích thích hợp dựa vào mục đích nghiên cứu và bản chất của dữ liệu. Do vậy, đầu tiên chúng ta tìm hiểu bản chất của dữ liệu thông qua khảo sát các cấp độ đo lường khác nhau vì mỗi cấp độ sẽ chỉ cho phép một số phương pháp nhất định mà thôi.

4.1. Khái niệm

- Số đo: là việc gán những dữ kiện lượng hoá hay những ký hiệu cho những hiện tượng quan sát. Chẳng hạn như những đặc điểm của khách hàng về sự chấp nhận, thái độ, thị hiếu hoặc những đặc điểm có liên quan khác đối với một sản phẩm mà họ tiêu dùng.

- Thang đo: là tạo ra một thang điểm để đánh giá đặc điểm của đối tượng nghiên cứu thể hiện qua sự đánh giá, nhận xét.

4.2. Các loại thang đo

- Thang đo danh nghĩa (Nominal scale):

Là loại thang đo sử dụng cho dữ liệu thuộc tính mà các biểu hiện của dữ liệu không có sự hơn kém, khác biệt về thứ bậc. Các con số không có mối quan hệ hơn kém, không thực hiện được các phép tính đại số. Các con số chỉ mang tính chất mã hoá. Ví dụ, tiêu thức giới tính ta có thể đánh số 1 là nam, 2 là nữ.

- Thang đo thứ bậc (Ordinal scale):

Là loại thang đo dùng cho các dữ liệu thuộc tính. Tuy nhiên trường hợp này biểu hiện của dữ liệu có sự so sánh. Ví dụ, trình độ thành thạo của công

nhân được phân chia ra các bậc thợ từ 1 đến 7. Phân loại giảng viên trong các trường đại học: Giáo sư, P.Giáo sư, Giảng viên chính, Giảng viên. Thang đo này cũng không thực hiện được các phép tính đại số.

- Thang đo khoảng (Interval scale):

Là loại thang đo dùng cho các dữ liệu số lượng. Là loại thang đo cũng có thể dùng để xếp hạng các đối tượng nghiên cứu nhưng khoảng cách bằng nhau trên thang đo đại diện cho khoảng cách bằng nhau trong đặc điểm của đối tượng. Với thang đo này ta có thể thực hiện các phép tính đại số trừ phép chia không có ý nghĩa. Ví dụ như điểm môn học của sinh viên. Sinh viên A có điểm thi là 8 điểm, sinh viên B có điểm là 4 thì không thể nói rằng sinh viên A giỏi gấp hai lần sinh viên B.

- Thang đo tỷ lệ (Ratio scale):

Là loại thang đo cũng có thể dùng dữ liệu số lượng. Trong các loại thang đo đây là loại thang đo cao nhất. Ngoài đặc tính của thang đo khoảng, phép chia có thể thực hiện được. Ví dụ, thu nhập trung bình 1 tháng của ông A là 2 triệu đồng và thu nhập của bà B là 4 triệu đồng, ta có thể nói rằng thu nhập trung bình trong một tháng của bà B gấp đôi thu nhập của ông A.

Tuỳ theo thang đo chúng ta có thể có một số phương pháp phân tích phù hợp, ta có thể tóm tắt như sau:

Phương pháp phân tích thống kê thích hợp với các thang đo

Loại thang đo	Đo lường độ tập trung	Đo lường độ phân tán	Đo lường tính tương quan	Kiểm định
1. Thang biểu danh	Mốt	Không có	Hệ số ngẫu nhiên	Kiểm định ²
2. Thang thứ tự	Trung vị	Số phần trăm	Dãy tương quan	Kiểm định dấu
3. Thang khoảng	Trung bình	Độ lệch chuẩn	Hệ số tương quan	Kiểm định t, F
4. Thang tỷ lệ	Trung bình tỷ lệ	Hệ số biến thiên	Tất cả các phép trên	Sử dụng tất cả các phép trên

4.3. Xác định nội dung thông tin, thu thập thông tin

Nói chung, tuỳ thuộc vào mục đích nghiên cứu để xác định những nội dung thông tin cần thu thập. Thông tin sử dụng cho quá trình nghiên cứu phải đảm bảo các yêu cầu cơ bản sau:

- Thích đáng: Số liệu thu thập phải phù hợp, đáp ứng được mục đích nghiên cứu. Số liệu đáp ứng được mục tiêu nghiên cứu có tính chất trực tiếp hoặc gián tiếp. Đối với những thông tin dễ tiếp cận thường thì ta sử dụng số

liệu trực tiếp, ví dụ muốn biết được nhu cầu của khách hàng chúng ta có thể hỏi trực tiếp khách hàng. Tuy nhiên, một số nội dung nghiên cứu mang tính chất nhạy cảm hoặc khó thu thập thì chúng ta có thể thu nhập những số liệu gián tiếp có liên quan, ví dụ để thu thập thu nhập của cá nhân chúng ta có thể thu thập những nội dung có liên quan như nghề nghiệp, đơn vị công tác, chức vụ, nhà ở, phương tiện đi lại...

- Chính xác: Các thông tin trong quá trình nghiên cứu phải có giá trị, đáng tin cậy

để các phân tích kết luận phản ánh được đặc điểm bản chất của hiện tượng.

- Kịp thời: Yêu cầu thông tin không những đáp ứng yêu cầu phù hợp, chính xác mà giá trị thông tin còn thể hiện ở chỗ nó có phục vụ kịp thời cho công tác quản lý và tiến trình ra các quyết định hay không.

- Khách quan: Tức là số liệu thu thập được không bị ảnh hưởng vào tính chủ quan của người thu thập cũng như người cung cấp số liệu và ngay cả trong thiết kế bảng câu hỏi. Yếu tố khách quan tưởng chừng thực hiện rất dễ dàng nhưng thực tế thì chúng ta khó có thể khắc phục vấn đề này một cách trọn vẹn, chúng ta chỉ có thể hạn chế yếu tố chủ quan một cách tối đa. Ví dụ chỉ cần một hành động đơn giản là tiếp cận với đáp viên là ít nhiều cũng ảnh hưởng đến kết quả trả lời của họ.

- Nguồn số liệu

Khi nghiên cứu một hiện tượng cụ thể, người nghiên cứu có thể sử dụng từ nguồn số liệu đã có sẵn đã được công bố hay chưa công bố hay tự mình thu thập các dữ liệu cần thiết cho nghiên cứu. Dựa vào cách thức này người ta chia dữ liệu thành 2 nguồn: dữ liệu thứ cấp và dữ liệu sơ cấp.

- + Dữ liệu thứ cấp (Secondary data):

Dữ liệu thứ cấp là các thông tin đã có sẵn và đã qua tổng hợp, xử lý. Loại dữ kiện này có thể thu thập từ các nguồn sau:

- (1) Số liệu nội bộ: là loại số liệu đã được ghi chép cập nhật trong đơn vị hoặc được thu thập từ các cuộc điều tra trước đây.

- (2) Số liệu từ các ấn phẩm của nhà nước: Các dữ liệu do các cơ quan thống kê nhà nước phát hành định kỳ như niên giám thống kê, các thông tin cập nhật hàng năm về tình hình dân số lao động, kết quả sản xuất của các ngành trong nền kinh tế, số liệu về văn hoá xã hội.

(3) Báo, tạp chí chuyên ngành: Các báo và tạp chí đề cập đến vấn đề có tính chất chuyên ngành như tạp chí thống kê, giá cả thị trường,...

(4) Thông tin của các tổ chức, hiệp hội nghề nghiệp: Viện nghiên cứu kinh tế, phòng thương mại

(5) Các công ty chuyên tổ chức thu thập thông tin, nghiên cứu và cung cấp thông tin theo yêu cầu.

Số liệu thứ cấp có ưu điểm là có thể chia sẻ chi phí, do đó nó có tính kinh tế hơn, số liệu được cung cấp kịp thời hơn. Tuy nhiên, dữ liệu thứ cấp thường là các thông tin cơ bản, số liệu đã được tổng hợp đã qua xử lý cho nên không đầy đủ hoặc không phù hợp cho quá trình nghiên cứu. Số liệu thứ cấp thường ít được sử dụng để dự báo trong thống kê, số liệu này thường được sử dụng trong trình bày tổng quan nội dung nghiên cứu, là cơ sở để phát hiện ra vấn đề nghiên cứu. Ngoài ra, số liệu thứ cấp còn được sử dụng để đối chiếu lại kết quả nghiên cứu để nhằm kiểm tra lại tính đúng đắn hoặc phát hiện ra những vấn đề mới để có hướng nghiên cứu tiếp.

+ Dữ liệu sơ cấp (Primary data): Là các thông tin thu thập từ các cuộc điều tra. Căn cứ vào phạm vi điều tra có thể chia thành 2 loại: Điều tra toàn bộ và điều tra chọn mẫu.

Điều tra toàn bộ: Là tiến hành thu thập thông tin trên tất cả các đơn vị thuộc tổng thể nghiên cứu.

Ưu điểm của điều tra toàn bộ là thu thập được thông tin về tất cả các đơn vị tổng thể. Tuy nhiên, loại điều tra này thường gặp phải một số trở ngại sau:

Số lượng đơn vị thuộc tổng thể chung thường rất lớn cho nên tiến hành điều tra toàn bộ mất nhiều thời gian và tốn kém.

Trong một số trường hợp do thời gian kéo dài dẫn đến số liệu kém chính xác do hiện tượng tự biến động qua thời gian.

Trong một số trường hợp điều tra toàn bộ sẽ không thực hiện được, ví dụ như kiểm tra chất lượng sản phẩm phải phá huỷ các đơn vị thuộc đối tượng nghiên cứu.

- Điều tra chọn mẫu: Để nghiên cứu tổng thể, ta chỉ cần lấy ra một số phần tử đại diện để nghiên cứu và từ đó suy ra kết quả cho tổng thể bằng các phương pháp thống kê.

Điều tra chọn mẫu thường được sử dụng vì các lý do sau:

- + Tiết kiệm chi phí
- + Cung cấp thông tin kịp thời cho quá trình nghiên cứu
- + Đáng tin cậy. Đây là yếu tố rất quan trọng, nó làm cho điều tra chọn mẫu trở nên có hiệu quả và được chấp nhận. Tuy nhiên, để có sự đáng tin cậy này chúng ta phải có phương pháp khoa học để đảm bảo tính chính xác để chỉ cần chọn ra một số quan sát mà có thể suy luận cho cả tổng thể rộng lớn – đó là nhờ vào các lý thuyết thống kê.

Việc sử dụng điều tra toàn bộ hay điều tra chọn mẫu phụ thuộc vào nhiều yếu tố có liên quan: kích thước tổng thể, thời gian nghiên cứu, khả năng về tài chính và nguồn lực, đặc điểm của nội dung nghiên cứu.

- Các phương pháp thu thập thông tin

Để thu thập dữ liệu ban đầu, tùy theo nguồn kinh phí và đặc điểm của đối tượng cần thu thập thông tin, ta có các phương pháp sau đây:

- + Quan sát: Là phương pháp thu thập dữ liệu bằng cách quan sát hành động, hành vi thái độ của đối tượng được điều tra. Ví dụ, nghiên cứu trẻ con yêu thích màu sắc nào, quan sát thái độ khách hàng khi dùng thử loại sản phẩm. Phương pháp này tỏ ra hiệu quả đối với các trường hợp đối tượng khó tiếp cận và tăng tính khách quan của đối tượng. Tuy nhiên, phương pháp này tỏ ra khá tốn kém nhưng lượng thông tin thu thập được ít.

- + Phương pháp gửi thư: Theo phương pháp này nhân viên điều tra gửi bằng câu hỏi đến đối tượng cung cấp thông tin qua đường bưu điện. Phương pháp gửi thư có thể thu thập thông tin với khối lượng lớn, tiết kiệm chi phí so với các phương pháp khác. Tuy nhiên tỷ lệ trả lời bằng phương pháp này tương đối thấp, đây là một nhược điểm rất lớn của phương pháp này.

- + Phỏng vấn bằng điện thoại: Phương pháp thu thập thông tin bằng cách phỏng vấn qua điện thoại. Phương pháp này thu thập được thông tin một cách nhanh chóng, tuy nhiên phương pháp này có nhược điểm: tốn kém, nội dung thu thập thông tin bị hạn chế.

- + Phỏng vấn trực tiếp:

Phương pháp phỏng vấn trực tiếp thích hợp cho những cuộc điều tra cần thu thập nhiều thông tin, nội dung của thông tin tương đối phức tạp cần thu thập một cách chi tiết. Phương pháp phỏng vấn trực tiếp cho 2 hình thức:

- (1) Phỏng vấn cá nhân. Nhân viên điều tra tiếp xúc với đối tượng cung cấp thông tin thường tại nhà riêng hoặc nơi làm việc. Thông thường phỏng vấn trực tiếp được áp dụng khi chúng ta cho tiến hành điều tra chính thức.

(2) Phỏng vấn nhóm. Nhân viên điều tra phỏng vấn từng nhóm để thảo luận về một vấn đề nào đó. Trường hợp này người ta thường sử dụng khi điều tra thử để kiểm tra lại nội dung của bảng câu hỏi được hoàn chỉnh chưa hoặc nhằm tìm hiểu một vấn đề phức tạp mà bản thân người nghiên cứu chưa nắm được một cách đầy đủ mà cần phải có ý kiến cụ thể từ những người am hiểu.

Sau đây ta có bảng tổng hợp một số ưu nhược điểm của các phương pháp thu thập thông tin.

Đặc điểm của các phương pháp thu thập thông tin

Tính chất	Phương pháp gợi thư	Phỏng vấn qua điện thoại	Phỏng vấn trực tiếp
Linh hoạt	Kém	Tốt	Tốt
Khối lượng thông tin	Đầy đủ	Hạn chế	Đầy đủ
Tốc độ thu thập thông tin	Chậm	Nhanh	Nhanh
Tỷ lệ câu hỏi được trả lời	Thấp	Cao	Cao
Chi phí	Tiết kiệm	Tốn kém	Tốn kém

Câu hỏi ôn tập:

1. Trình bày thống kê là gì? Các khái niệm thường dùng trong thống kê?
2. Trình bày các loại thang đo?

CHƯƠNG 2

QUÁ TRÌNH THỐNG KÊ

Mã chương: MH10_CH02

Giới thiệu:

Giới thiệu về các khái niệm như phân tổ thống kê, các loại phân tổ thống kê. Xác định, phân tổ thống kê và các bảng thống kê.

Mục tiêu:

- Kiến thức: Trình bày được khái niệm phân tổ thống kê, nguyên tắc phân tổ thống kê, các tiêu thức, bảng thống kê, đồ thị thống kê.

- Kỹ năng: Áp dụng được phương pháp phân tổ thống kê trong nghiên cứu một số vấn đề xã hội và lập được bảng thống kê, đồ thị thống kê trong một số trường hợp.

- Thái độ: Cẩn thận, tỉ mỉ, chính xác, sáng tạo.

Nội dung chính:

1. Phân tổ thống kê

1.1. Khái niệm

Phân tổ còn được gọi là phân lớp thống kê là căn cứ vào một hay một số tiêu thức để chia các đơn vị tổng thể ra thành nhiều tổ (lớp, nhóm) có tính chất khác nhau.

1.2. Nguyên tắc phân tổ

Một cách tổng quát tổng thể phải được phân chia một cách trọn vẹn, tức là một đơn vị của tổng thể chỉ thuộc một tổ duy nhất và một đơn vị thuộc một tổ nào đó phải thuộc tổng thể.

1.3. Phân tổ theo tiêu thức thuộc tính

- Trường hợp tiêu thức thuộc tính chỉ có một vài biểu hiện thì mỗi biểu hiện của tiêu thức thuộc tính có thể chia thành một tổ. Ví dụ, tiêu thức giới tính.

- Trường hợp tiêu thức thuộc tính có nhiều biểu hiện, ta ghép nhiều nhóm nhỏ lại với nhau theo nguyên tắc các nhóm ghép lại với nhau có tính chất giống nhau hoặc gần giống nhau. Ví dụ phân tổ trong công nghiệp chế biến: Thực phẩm và đồ uống, thuốc lá, dệt,...

1.4. Phân tổ theo tiêu thức số lượng

- Trường hợp tiêu thức số lượng có ít biểu hiện, thì cứ mỗi một lượng biến có thể thành lập một tổ.

Ví dụ 2.1: phân tổ công nhân trong một xí nghiệp dệt theo số máy do mỗi công nhân thực hiện.

Số máy/Công nhân	Số công nhân
10	3
11	7
12	20
13	50
14	35
15	15

- Trường hợp tiêu thức số lượng có nhiều biểu hiện, ta phân tổ khoảng cách mỗi tổ và mỗi tổ có một giới hạn:

- Giới hạn trên: lượng biến nhỏ nhất của tổ.
- Giới hạn dưới: lượng biến lớn nhất của tổ.

Tuỳ theo mục đích nghiên cứu, người ta phân ra 2 loại phân tổ đều và phân tổ không đều.

- Phân tổ đều: Là phân tổ có khoảng cách tổ bằng nhau. Thông thường nếu chỉ vì mục đích nghiên cứu phân phối của tổng thể hoặc làm cho bảng thống kê gọn lại thì ta thường dùng phương pháp này.

- Để xác định số tổ hình như không có một tiêu chuẩn tối ưu nó phụ thuộc vào kinh nghiệm. Dưới đây là một cách phân chia tổ mang tính chất tham khảo.

- Xác định số tổ (Number off classes):

$$\text{Số tổ} = (2 \times n)^{0,3333} \quad n: \text{Số đơn vị tổng thể}$$

- Xác định khoảng cách tổ (Class interval):

$$k = \frac{X_{\max} - X_{\min}}{\text{Số tổ}}$$

So to

- Xác định tần số (Frequency) của mỗi tổ: bằng cách đếm các quan sát rơi vào giới hạn của tổ đó.

Một số qui ước khi lập bảng phân tổ:

- Trường hợp phân tổ theo tiêu thức số lượng rời rạc thì giới hạn trên và giới hạn dưới của 2 tổ kế tiếp nhau không được trùng nhau.

Ví dụ 2.2: Các xí nghiệp ở tỉnh X được phân tổ theo tiêu số lượng công nhân:

Số lượng công nhân	Số xí nghiệp
≤ 100	80
101 – 200	60
201 – 500	6
501 – 1.000	4
1.001 – 2.000	1
Tổng	151

- Trường hợp phân tổ theo tiêu thức số lượng loại liên tục, thường có qui ước sau:

* Giới hạn trên và giới hạn dưới của 2 tổ kế tiếp trùng nhau.

* Quan sát có lượng biến bằng đúng giới hạn trên của một tổ nào đó thì đơn vị đó được xếp vào tổ kế tiếp.

Ví dụ 2.3: phân tổ các tổ chức thương nghiệp theo doanh thu.

Doanh thu (triệu đồng)	Số tổ chức thương nghiệp
≤ 1.000	2
1.000-2.000	9
2.000-3.000	12
3.000-4.000	7
Tổng	30

1.5. Bảng phân phối tần số (Frequency table)

Sau khi phân tổ chúng ta có thể trình bày số liệu bằng cách sử dụng bảng phân phối tần số để biết được một số tính chất cơ bản của hiện tượng nghiên cứu.

Lượng biến	Tần số	Tần số tương đối	Tần số tích lũy
------------	--------	------------------	-----------------

x_1	f_1	f_1/n	f_1
x_2	f_2	$\dots f_i/n$	$f_1 + f_2$
$\dots x_i$	$\dots f_i$	$\dots$	$\dots$
$\dots$	$\dots$	f_k/n	$f_1 + f_2 + \dots + f_i$
x_k	f_k		$\dots$
Cộng	$\sum_{i=1}^k f_i = n$	1	

Trong đó lượng biến có thể là giá trị cụ thể hoặc là một khoảng.

1.6. Các loại phân tổ thống kê

- Phân tổ kết cấu:

Trong công tác nghiên cứu thống kê, các bảng phân tổ kết cấu được sử dụng rất phổ biến nhằm mục đích nêu lên bản chất của hiện tượng trong điều kiện nhất định và để nghiên cứu xu hướng phát triển của hiện tượng qua thời gian.

Ví dụ 2.4: Để xem xét cơ cấu giữa các nhóm ngành trong một quốc gia nào đó ta lập bảng như sau:

Bảng 2.1. Cơ cấu tổng sản phẩm của quốc gia X theo nhóm ngành, 2003 -2007

Đơn vị tính: %.

Tổng sản phẩm theo nhóm ngành	2003	2004	2005	2006	2007
Nông, lâm nghiệp và thủy sản	24,53	23,24	23,03	22,54	21,76
Công nghiệp và xây dựng	36,73	38,13	38,49	39,47	40,09
Dịch vụ	38,74	38,63	38,48	37,99	38,15
Tổng	100,00	100,00	100,00	100,00	100,00

Qua bảng kết cấu trên, ta thấy có thấy sự thay đổi về dịch chuyển cơ cấu ngành: Nhóm ngành công nghiệp và xây dựng có xu hướng tăng, nhóm ngành nông, lâm, thủy sản có xu hướng giảm,...

- Phân tổ liên hệ:

Khi tiến hành phân tổ liên hệ, các tiêu thức có liên hệ với nhau được phân biệt thành 2 loại tiêu thức nguyên nhân và tiêu thức kết quả. Phân tổ liên hệ có thể được vận dụng để nghiên cứu mối liên hệ giữa nhiều tiêu thức: mối liên hệ giữa năng suất với lượng phân bón, nghiên cứu giữa năng suất lao động của công nhân với tuổi nghề, bậc thợ, trình độ trang bị kỹ thuật,...

Ví dụ 2.5: Ta có bảng phân tổ liên hệ sau:

Bảng 2.2. Liên hệ giữa năng suất lao động với trình độ kỹ thuật nghề nghiệp của quốc gia X năm 2007

Trình độ kỹ thuật	Tuổi nghề (Năm)	Số công nhân	Sản lượng cả năm (tấn)	Năng suất lao động bình quân
Đã được đào tạo kỹ thuật	dưới 5	15	1.125	75
	5-10	40	3.750	94
	10-15	40	4.200	105
	15-20 trên	15	1.725	115
	20	10	1.200	120
Cả tổ	-	120	12.000	100
Chưa được đào tạo kỹ thuật	dưới 5	10	510	51
	5-10	30	2.140	71
	10-15	20	1.540	79
	15-20 trên	10	860	86
	20	10	910	91
Cả tổ	-	80	6.000	75
Chung cho cả doanh nghiệp	-	200	18.000	90

2. Bảng thống kê (Statistical table)

Sau khi tổng hợp các tài liệu điều tra thống kê, muốn phát huy tác dụng của nó đối với phân tích thống kê, cần thiết phải trình bày kết quả tổng hợp theo một hình thức thuận lợi nhất cho việc sử dụng sau này.

2.1. Khái niệm

Bảng thống kê là một hình thức trình bày các tài liệu thống kê một cách có hệ thống, hợp lý và rõ ràng, nhằm nêu lên các đặc trưng về mặt lượng của hiện tượng nghiên cứu. Đặc điểm chung của tất cả các bảng thống kê là bao giờ cũng có những con số của từng bộ phận và có mối liên hệ mật thiết với nhau.

2.2. Cấu thành bảng thống kê

- Về hình thức: Bảng thống kê bao gồm các hàng, cột, các tiêu đề, tiêu mục và các con số.

Các hàng cột thể hiện qui mô của bảng, số hàng và cột càng nhiều thì bảng thống kê càng lớn và càng phức tạp.