

NGUYÊN LÝ THỐNG KÊ KINH TẾ (867001, 45 Tiết)

GV: Hoàng Đức Thắng
(hdthang@sgu.edu.vn)

Khoa Toán-Ứng dụng
Đại học Sài Gòn

2019

Nội dung môn học:

- Chương 1: Những vấn đề lý luận chung của thống kê.
- Chương 2: Quá trình nghiên cứu thống kê.
- Chương 3: Thống kê các mức độ của hiện tượng nghiên cứu.
- Chương 4: Điều tra chọn mẫu.
- Chương 5: Tương quan hồi quy.
- Chương 6: Dây số thời gian.
- Chương 7: Chỉ số thống kê.

Tài liệu tham khảo:

- [1] Mai Văn Nam, Giáo trình nguyên lý thống kê kinh tế, Nhà xuất bản Văn hóa thông tin, (2008).
- [2] Trần Đình Kỳ, Nguyễn Văn Phúc, Giáo trình nguyên lý thống kê, Nhà xuất bản lao động, (2012).
- [3] Hà Văn Sơn, Giáo trình nguyên lý thống kê kinh tế, Trường Đại học Kinh tế, Nhà xuất bản Thống kê, (2010).

Kiểm tra, đánh giá môn học

- Điểm chuyên cần, hệ số 0.1:
- Đi học đầy đủ: 10đ
 - Nghỉ 1 buổi, hoặc đi trễ 2 buổi: - 1đ
 - Được vắng 3 buổi, từ buổi thứ 4 mỗi buổi vắng trừ 1 điểm.
- Điểm kiểm tra giữa kỳ, hệ số 0.3:
- Tự luận, không dùng tài liệu.
- Điểm kiểm tra cuối kỳ, hệ số 0.6:
- Tự luận, không dùng tài liệu.
- Điểm cộng:
- Lên bảng 1 lần, làm đúng: + 0,5đ.
 - Được cộng tối đa: 2đ.

Chương 1: NHỮNG VẤN ĐỀ LÝ LUẬN CHUNG CỦA THỐNG KÊ

1. Khái niệm về thống kê

- ▷ **Thống kê (Statistics)** là môn khoa học bao gồm các phương pháp thu thập, phân tích, trình bày, diễn giải dữ liệu.
- ▷ Thống kê được chia làm hai lĩnh vực:
 - Thống kê mô tả** bao gồm các phương pháp thu thập, trình bày dữ liệu, và tính toán các đặc trưng nhằm mô tả đối tượng nghiên cứu.
 - Thống kê suy diễn** bao gồm các phương pháp ước lượng, kiểm định, phân tích mối liên hệ, dự đoán ... trên cơ sở các thông tin thu thập từ mẫu từ đó cho ta những hiểu biết về tổng thể.
- ▷ **Đối tượng nghiên cứu của thống kê học** là mặt lượng trong mối liên hệ mật thiết với mặt chất của các hiện tượng số lớn, trong điều kiện thời gian và địa điểm cụ thể.

2. Một số khái niệm của thống kê

2.1. Tổng thể

▷ **Tổng thể (Populations)** là tập hợp tất cả các phần tử (hay đơn vị) mà ta quan tâm trong một nghiên cứu cụ thể.

Ví dụ 1

Muốn tính thu nhập trung bình của một hộ gia đình ở Thành phố Hồ Chí Minh, thì

- tổng thể là tất cả các hộ gia đình ở Thành phố Hồ Chí Minh,
- mỗi phần tử của tổng thể là một hộ gia đình ở Thành phố.

Căn cứ vào sự nhận biết các đơn vị trong tổng thể, ta có thể chia thành hai loại: *tổng thể bộc lộ* và *tổng thể tiềm ẩn*.

Căn cứ vào mục đích nghiên cứu, ta có thể chia làm hai loại tổng thể: *tổng thể đồng chất* và *tổng thể không đồng chất*.

Căn cứ vào phạm vi nghiên cứu, người ta còn phân biệt: *tổng thể chung* và *tổng thể bộ phận*.

▷ **Tổng thể bộc lộ** là tổng thể có các phần tử được biểu hiện một cách rõ ràng, dễ xác định.

Ví dụ 2

1. Tổng thể sinh viên học môn NLTKKT, Trường Đại học Sài Gòn học kì III năm 2018-2019.
2. Tổng thể xe máy được cấp đăng ký trong tháng 8 năm 2019 tại TP HCM.

▷ **Tổng thể tiềm ẩn** là tổng thể mà các phần tử của nó không được biểu hiện rõ ràng, khó xác định.

Ví dụ 3

1. Tổng thể những người hâm mộ bóng đá.
2. Tổng thể những người bị nhiễm virus lao.

▷ **Tổng thể đồng chất** bao gồm những phần tử cùng chung những đặc điểm chủ yếu liên quan đến mục đích nghiên cứu.

▷ **Tổng thể không đồng chất** bao gồm những phần tử khác nhau về loại hình, khác nhau về những đặc điểm chủ yếu có liên quan đến mục đích nghiên cứu.

Ví dụ 4

Khi nghiên cứu về thu nhập trung bình của sinh viên trường Đại học Sài Gòn mới ra trường:

- Nếu xét tổng thể sinh viên khối ngành kinh tế \Rightarrow tổng thể đồng chất.
- Nếu xét tổng thể tất cả sinh viên \Rightarrow tổng thể không đồng chất.

Việc xác định tổng thể đồng chất hay không đồng chất phụ thuộc vào mục đích nghiên cứu. Những kết luận rút ra từ nghiên cứu thống kê chỉ có ý nghĩa khi nghiên cứu một tổng thể đồng chất.

▷ **Tổng thể chung** bao gồm tất cả các phần tử của hiện tượng nghiên cứu.

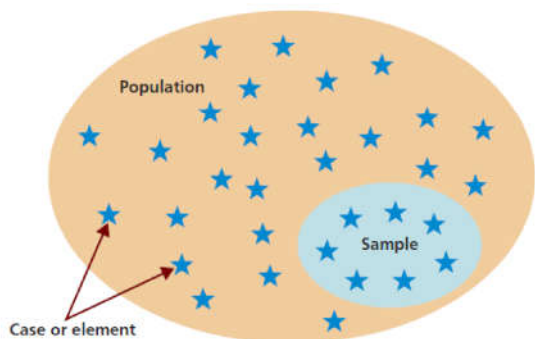
▷ **Tổng thể bộ phận** chỉ chứa một bộ phận của tổng thể chung.

Ví dụ 5

1. Tổng dân số Việt Nam là tổng thể chung.
2. Tổng dân số nữ của Việt Nam là tổng thể bộ phận.

2.2. Mẫu

▷ **Mẫu (Samples)** là một tập con của tổng thể mà ta lấy ra để nghiên cứu theo một phương pháp lấy mẫu nào đó. Các đặc trưng mẫu được sử dụng để suy rộng ra các đặc trưng của tổng thể.



2.3. Biến (tiêu thức)

▷ **Biến** là khái niệm chỉ các đặc điểm được quan tâm của phần tử tổng thể.

Trong thống kê, biến được phân biệt thành hai loại:

▷ **Biến định tính:** là tính chất của phần tử tổng thể được khảo sát, không có biểu hiện trực tiếp bằng số.

Ví dụ 6

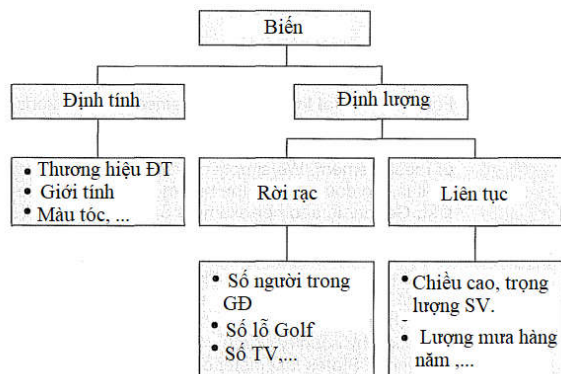
Giới tính, trình độ văn hóa, nghề nghiệp, dân tộc, tôn giáo, loại hình doanh nghiệp, ...

▷ **Biến định lượng:** là đặc điểm của phần tử tổng thể có biểu hiện trực tiếp bằng con số.

Ví dụ 7

Tuổi, chiều cao, trọng lượng, năng suất làm việc, ...

- ▷ **Biến định lượng rời rạc:** là biến có giá trị hữu hạn, hoặc vô hạn đếm được.
- ▷ **Biến định lượng liên tục:** là biến có giá trị có thể lấp kín cả một khoảng trên trục số.



- ▷ **Quan sát (Observation):** là tập hợp các số đo thu được của một phần tử.

2.4. Chỉ tiêu thống kê

Chỉ tiêu thống kê phản ánh đặc điểm về quy mô, tốc độ phát triển, cơ cấu, trình độ phổ biến, quan hệ tỷ lệ của bộ phận hoặc toàn bộ hiện tượng kinh tế - xã hội trong điều kiện không gian và thời gian cụ thể. Chỉ tiêu thống kê gồm tên chỉ tiêu và trị số của chỉ tiêu. (Luật thống kê 2015)

Ví dụ 8

Theo niên giám thống kê năm 2010, tổng sản phẩm trong nước (GDP) của Việt Nam năm 2010 (tính sơ bộ) theo giá thực tế là 1 980 914 tỷ đồng; số dân trung bình của Việt Nam năm 2010 là 86 927,7 nghìn người và mật độ dân số cả nước là 263 người/Km².

- Chỉ tiêu: số dân trung bình của Việt Nam năm 2010.
- Trị số: 86 927,7 (đơn vị: nghìn người).

- ▷ **Chỉ tiêu khối lượng:** phản ánh quy mô, khối lượng của hiện tượng nghiên cứu theo thời gian và địa điểm cụ thể.

Ví dụ 9

- Số dân Việt Nam vào 0 giờ ngày 1/4/2009 là 85 846 997 người.
- Số doanh nghiệp, số vốn cố định, số người sử dụng app gọi xe ...
- Tổng sản phẩm quốc nội (GDP).

- ▷ **Chỉ tiêu chất lượng:** Phản ánh các đặc điểm về mặt chất của hiện tượng như trình độ phổ biến, mức độ tốt xấu và quan hệ so sánh trong tổng thể.

Ví dụ 10

Giá thành đơn vị sản phẩm là chỉ tiêu chất lượng, nó biểu hiện quan hệ so sánh giữa tổng giá thành và số lượng sản phẩm sản xuất ra, đồng thời nó phản ánh tính chất phổ biến về mức chi phí cho một sản phẩm đã được sản xuất.

3. Các loại thang đo thường dùng trong thống kê

- ▷ **Thang đo định danh (Nominal scale)** là loại thang đo dùng cho các biến định tính. Người ta thường sử dụng các mã số (code) để phân loại các đối tượng, ngoài vai trò này các mã số không mang ý nghĩa nào khác.

Ví dụ 11

Giới tính: 1. Nam 2. Nữ.

Tình trạng hôn nhân:

1. Độc thân 2. Có gia đình 3. Ly dị 4. Trường hợp khác.

Sinh viên khoa:

1. QTKT 2. TCKT 3. Khoa khác

- ▷ **Thang đo thứ bậc (Ordinal scale)** thường dùng cho các biến định tính (đôi khi cũng áp dụng cho các biến định lượng). Trong thang đo thứ bậc, giữa các biểu hiện của biến có quan hệ thứ bậc hơn kém, sự chênh lệch giữa các biểu hiện không bằng nhau, vì vậy không thực hiện được các phép toán (+, -, :, :)

Ví dụ 12

Trình độ thành thạo của người công nhân được chia là các bậc thợ: 1, 2, 3, 4, ...

Chất lượng sản phẩm được chia thành: SP loại 1, SP loại 2, SP loại 3.

Các con số trên thể hiện trật tự hơn kém của các biểu hiện, tuy nhiên không nhất thiết theo logic toán học.

Chất lượng SP loại 1 tốt nhất, nhưng bậc thợ càng cao tay nghề càng giỏi.

➤ **Thang đo khoảng** thường dùng cho biến định lượng (đôi khi cũng được áp dụng cho biến định tính). Thang đo khoảng là thang đo thứ bậc có các khoảng cách đều nhau nhưng không có điểm gốc là 0. Ta có thể thực hiện phép toán (+, -)

➤ **Thang đo tỉ lệ** dùng cho biến định lượng. Thang đo tỉ lệ có đầy đủ các đặc tính của thang đo khoảng tức là có thể áp dụng phép tính cộng trừ, ngoài ra nó còn có trị số 0 "thật". Đây là loại thang đo định lượng chặt chẽ nhất.

Ví dụ 13

1. Số tiền 10 Tr.đ sẽ cao gấp đôi số tiền 5 Tr.đ ⇒ thang đo tỉ lệ.
Nếu ta đổi sang USD thì nó vẫn gấp đôi.
Nếu ta bị mất thì ta có 0 đồng.

2. Nhiệt độ C không phải là thang đo tỉ lệ, vì không thể nói 12^oC (53,6^oF) gấp đôi 6^oC (42,8^oF) (10^oC và -10^oC thì sao?)
Nhiệt độ 0^o chỉ là quy ước nhiệt độ mà tại đó nước đóng băng.

Biến định lượng	Thang đo
Nhiệt độ	khoảng
Thời gian	khoảng
Giá cổ phiếu	tỉ lệ
Tuổi	tỉ lệ
Chiều cao	tỉ lệ
Điểm thi SAT	tỉ lệ

Theo bảng trên, ta có

- Mỗi cổ phiếu của 1 công ty là một *phần tử*.
Bảng trên có 10 phần tử.
- Có 5 *biến*: TTCK, Mã CP, Xếp hạng BW, Giá CP, Lợi nhuận.
- *Quan sát* công ty AB là bộ: (N, ABT, 90, 46, 2,02).
Bảng trên gồm 10 quan sát.
- *Thang đo danh nghĩa* dùng cho biến: TTCK, Mã CP.
- *Thang đo thứ bậc* dùng cho biến: Xếp hạng của BW.
- *Thang đo tỉ lệ* dùng cho biến: Giá cổ phiếu, lợi nhuận trên cổ phần.

Ví dụ 14

Biến đ.tính	Thang đo định danh				
Giới tính	Nam		Nữ		
Có xe hơi	có		không		
Trình độ	đại học	trung học	tiểu học	mù chữ	
Lĩnh vực	Sản xuất	Xây dựng	Dịch vụ	Khác	
	Thang đo thứ tự				
Xếp hạng SV	A	B	C	D	E
Đánh giá KS	5*	4*	3*	2*	1*
Đánh giá SP	tốt		khá		tệ
Thu nhập	< 5 Tr.đ		5-10 Tr.đ		> 10 Tr.đ

Ví dụ 15

Cho bảng thông tin của 10 CT nằm trong danh sách S&P 500.

	TT chứng khoán	Mã cổ phiếu	Xếp hạng BW	Giá cổ phiếu	lợi nhuận /cổ phần
Tên công ty					
Abbott Laboratoriess	N	ABT	90	46	2,02
Altria Gruop	N	MO	148	66	4,57
Apollo Gruop	NQ	APOL	174	74	0,90
Bank of New York	N	BK	305	30	1,85
Bristol Myers Squibb	N	BMY	346	26	1,21
Cincinnati Financial	NQ	CINF	161	45	2,73
Comcast	NQ	CMCSA	296	32	0,43
Deere	N	DE	36	71	5,77
Ebay	NQ	EBAY	19	43	0,57
Federated Dept. Stores	N	FD	353	56	3,86

Chương 1

Những vấn đề lý luận chung của thống kê

Bài tập 1.1. Một phiếu đăng ký hoàn thuế thu nhập cá nhân hỏi về các thông tin sau: thu nhập, số nhân khẩu phụ thuộc, tình trạng hôn nhân, mức nộp thuế.

Hãy xác định các loại thang đo và loại biến.

Bài tập 1.2. Năm mùi vị kem được xếp hạng theo sở thích, cần sử dụng loại thang đo nào? Thang đo đối với màu sắc trong môn Karate là gì?

Bài tập 1.3. Công ty điện lực TP HCM xây dựng bảng câu hỏi nhằm nghiên cứu khách hàng về các yếu tố sau:

1. Giới tính chủ hộ.
2. Tuổi chủ hộ.
3. Nghề nghiệp chủ hộ.
4. Số thành viên trong hộ.
5. Có sử dụng máy nước nóng năng lượng mặt trời? (Có / Không).
6. Số máy điều hòa nhiệt độ sử dụng trung bình mỗi ngày (qui đổi thành máy 1HP).
7. Thời gian trung bình sử dụng máy điều hòa mỗi ngày.
8. Thu nhập/tháng (triệu đồng) của chủ hộ.
 < 5 $5 - 7$ $7 - 10$ ≥ 10
9. Tổng thu nhập mỗi tháng của hộ (triệu đồng).
10. Số tiền điện trung bình trong hóa đơn mỗi tháng.

Hãy xác định các loại thang đo và loại biến (định tính, định lượng, rời rạc, liên tục.).

Bài tập 1.4. (HKI - 2017- 2018)

Một nghiên cứu xây dựng mô hình biến động giá bán một căn nhà (tỷ đồng) phụ thuộc vào diện tích căn nhà (m^2) và tình trạng hợp pháp của căn nhà đó. Hãy cho biết các biến trên thuộc loại biến nào (định tính hay định lượng) và sử dụng thang đo thích hợp để đo lường các biến đó.

Bài tập 1.5. (HKII - 2017- 2018)

Thực hiện nghiên cứu so sánh mức lương trung bình của một sinh viên vừa tốt nghiệp từ các ngành học khác nhau tại một trường đại học X có là như nhau hay không theo anh chị, ta nên phân tổ theo tiêu thức nào? Hãy cho biết các biến trên thuộc loại biến nào?

Bài tập 1.6. (HKIII - 2017- 2018)

Một công ty định phí bảo hiểm muốn xây dựng mô hình để dự đoán tuổi thọ của một cá nhân. Công ty thu thập dữ liệu về tuổi thọ, số giờ tập thể dục và lượng cholesterol (mg) của 50 khách hàng. Hãy cho biết các biến trên thuộc loại biến nào (định tính hay định lượng) và sử dụng thang đo thích hợp để đo lường các biến đó.

Bài tập 1.7. Tạp chí Condé Nast Traveler thực hiện một cuộc khảo sát hàng năm người đọc để xác định nơi lưu trú tốt nhất trên toàn thế giới. Bảng 1.6 cho thấy mẫu 9 khách sạn ở Châu Âu. Giá phòng đôi chuẩn trong mùa du lịch của khách sạn này từ giá thấp nhất là \$ đến giá cao nhất là \$\$\$\$\$. Điểm đánh giá chung của người đọc về từng khách sạn bao gồm phòng, dịch vụ, nhà hàng, vị trí, không khí, và khu vực công cộng. Điểm đánh giá chung càng cao thì mức độ hài lòng càng cao.

1. Có bao nhiêu phần tử trong tập dữ liệu?
2. Có bao nhiêu biến trong tập dữ liệu?
3. Biến nào là biến định tính, định lượng?
4. Thang đo của từng biến này là gì?

Bài tập 1.8. Khách sạn Ritz-Carlton sử dụng một bảng thăm dò ý kiến khách hàng để thu thập dữ liệu hiệu quả về các dịch vụ ăn uống và giải trí. Khách hàng được yêu cầu đánh giá sáu yếu tố sau:

Chào mừng, dịch vụ, thức ăn, sự hấp dẫn của thực đơn, bầu không khí, đánh giá chung.

Dữ liệu được ghi lại cho mỗi yếu tố với

1: tạm; 2: trung bình; 3: tốt; 4: xuất sắc.

Trả lời của khách hàng cung cấp dữ liệu cho sáu biến.

1. Các biến này là định tính hay định lượng.
2. Loại thang đo cho từng biến?

Bài tập 1.9. Hệ thống âm thanh tất cả-trong-một, được gọi là minisystems, thường bao gồm AM / FM, một đầu hát băng cassette, và một đầu CD nhiều đĩa có kích thước cỡ quyển sách với hai loa riêng biệt, Dữ liệu trong bảng 1.7 cho thấy giá bán lẻ, chất lượng âm thanh, khả năng chứa đĩa CD, độ nhạy điều chỉnh, chọn lọc FM, và số lượng các hộp băng của một mẫu 10 minisystems. (Consumer Reports Buying Guide 2002).

1. Tập dữ liệu này chứa bao nhiêu phần tử?
2. Tổng thể này là gì?

3. Có bao nhiêu biến trong tập dữ liệu này? biến bào là biến định tính, định lượng?

TABLE 1.6 RATINGS FOR NINE PLACES TO STAY IN EUROPE

Name of Property	Country	Room Rate	Number of Rooms	Overall Score
Graveteye Manor	England	\$\$	18	83.6
Villa d'Este	Italy	\$\$\$\$	166	86.3
Hotel Prem	Germany	\$	54	77.8
Hotel d'Europe	France	\$\$	47	76.8
Palace Luzern	Switzerland	\$\$	326	80.9
Royal Crescent Hotel	England	\$\$\$	45	73.7
Hotel Sacher	Austria	\$\$\$	120	85.5
Duc de Bourgogne	Belgium	\$	10	76.9
Villa Gallici	France	\$\$	22	90.6

Source: Condé Nast Traveler, January 2000.

TABLE 1.7 A SAMPLE OF 10 MINISYSTEMS

Brand and Model	Price (\$)	Sound Quality	CD Capacity	FM Tuning	Tape Decks
Aiwa NSX-AJ800	250	Good	3	Fair	2
JVC FS-SD1000	500	Good	1	Very Good	0
JVC MX-G50	200	Very Good	3	Excellent	2
Panasonic SC-PM11	170	Fair	5	Very Good	1
RCA RS 1283	170	Good	3	Poor	0
Sharp CD-BA2600	150	Good	3	Good	2
Sony CHC-CL1	300	Very Good	3	Very Good	1
Sony MHC-NX1	500	Good	5	Excellent	2
Yamaha GX-505	400	Very Good	3	Excellent	1
Yamaha MCR-E100	500	Very Good	1	Excellent	0