XEM XÉT GIẢNG DẠY KIẾN THỨC THỐNG KÊ TRONG CHƯƠNG TRÌNH GIÁO DỤC PHỔ THÔNG QUA BỘ SÁCH TOÁN CÁNH DIỀU

TS. Phạm Đăng Quyết[*]

Tóm tắt:

Giáo dục thống kê trong trường phổ thông là chủ đề hiện được quan tâm rất lớn trong bối cảnh Việt Nam đang tiến hành đổi mới nền giáo dục. Bài viết này tập trung xem xét nội dung kiến thức thống kê cấp trung học cơ sở (từ lớp 6 đến lớp 9) qua bộ sách Toán Cánh diều và so sánh, đối chiếu nội dung thống kê có trong Chương trình giáo dục phổ thông môn toán 2018 với nội dung thống kê môn Toán của Khan Academy.

Qua bài viết này chúng tôi muốn chỉ ra những bất cập về nội dung kiến thức thống kê có trong Chương trình giáo dục phổ thông môn toán 2018, mặc dù chúng được xây dựng trên cơ sở lý luận và thực tiễn giáo dục phổ thông, do các nhà khoa học giáo dục - những người am hiểu giáo dục phổ thông môn toán biên soạn. Đồng thời cũng chỉ ra những thiếu sót và hạn chế trong nội dung kiến thức thống kê có trong bộ sách giáo khoa Toán cấp trung học cơ sở.

Hy vọng những ý kiến cá nhân này có thể góp phần làm tư liệu cho các nhà chuyên môn tham khảo khi biên soạn và điều chỉnh chương trình và sách giáo khoa môn toán trong những năm tiếp theo.

Từ khóa: Chương trình giáo dục phổ thông năm 2018, sách giáo khoa Toán phổ thông, thống kê và xác suất, kiến thức thống kê.

1. Đặt vấn đề

Giáo dục thống kê trong trường phổ thông là chủ đề hiện được quan tâm rất lớn trong bối cảnh Việt Nam đang tiến hành đổi mới nền giáo dục. Hiện nay kiến thức thống kê và xác suất đã được giảng dạy trong chương trình giáo dục phổ thông thông qua ba bộ sách giáo khoa Toán Cánh diều (Nhà XB Đại học Sư phạm), Chân trời sáng tạo (Nhà XB Giáo dục Việt Nam) và Kết nối tri thức với cuộc sống (Nhà XB Giáo dục Việt Nam). Nội dung kiến thức môn học phải đáp ứng yêu cầu cần đạt có trong Khung Chương trình giáo dục phổ thông môn toán theo chương trình mới của Bộ Giáo dục và Đào tạo (năm 2018). Chương trình giáo dục phổ thông môn toán xác định mục tiêu đối với từng cấp học như sau:

Thống kê và Xác suất cấp tiểu học: Một số yếu tố thống kê và xác suất đơn giản; giải quyết một số vấn đề thực tiễn đơn giản gắn với một số yếu tố thống kê và xác suất.

Thống kê và Xác suất cấp trung học cơ sở: Thu thập, phân loại, biểu diễn, phân tích và xử lí dữ liệu thống kê; phân tích dữ liệu thống kê thông qua tần số, tần số tương đối; nhận biết một số quy luật thống kê đơn giản trong thực tiễn; sử dụng thống kê để hiểu các khái niệm cơ bản về xác suất thực nghiệm của một biến cố và xác suất của một biến cố; nhận biết ý nghĩa của xác suất trong thực tiễn.

[*] *Hội Thống kê Việt Nam*

Thống kê và Xác suất cấp trung học phổ thông: Hoàn thiện khả năng thu thập, phân loại, biểu diễn, phân tích và xử lí dữ liệu thống kê; sử dụng các công cụ phân tích dữ liệu thống kê thông qua các số đặc trưng đo xu thế trung tâm và đo mức độ phân tán cho mẫu số liệu không ghép nhóm và ghép nhóm; sử dụng các quy luật thống kê trong thực tiễn; nhận biết các mô hình ngẫu nhiên, các khái niệm cơ bản của xác suất và ý nghĩa của xác suất trong thực tiễn.

Do đó, thay vì xem xét nội dung kiến thức môn học Thống kê trong cả ba bộ sách Toán, chúng tôi sẽ xem xét kiến thức thống kê được giảng dạy trong một bộ sách giáo khoa Toán là sách Toán Cánh diều.

2. Nội dung nghiên cứu

Kiến thức thống kê bắt đầu được giảng dạy trong chương trình giáo dục phổ thông bắt đầu từ lớp 2 đến lớp 11. Ở lớp 2 kiến thức thống kê về thu thập, kiểm đếm và biểu đồ tranh được lồng ghép vào Mục 4. Phép cộng, phép trừ trong phạm vi 1000. Ở lớp 3 kiến thức thống kê về thu thập, phân loại, ghi chép số liệu thống kê và bảng số liệu thống kê được lồng ghép vào Mục 4. Cộng, trừ, nhân, chia trong phạm vi 100000. Ở lớp 4 kiến thức thống kê về dãy số liệu thống kê, biểu đồ cột được lồng ghép vào Mục IV. Các phép tính với phân số.

Bắt đầu từ lớp 5 kiến thức thống kê và xác suất mới được giảng dạy thành một chuyên mục riêng "Thống kê và xác suất" (lớp 5) hay một Chương riêng "Một số yếu tố thống kê và xác suất" (lớp 6 đến lớp 11). Lớp 12 không giảng dạy thống kê mà chỉ giảng dạy xác suất với tiêu đề Chương là "Một số yếu tố xác suất".

Bài viết này tập trung xem xét nội dung kiến thức thống kê cấp trung học cơ sở (từ lớp 6 đến lớp 9) và so sánh, đối chiếu nội dung thống kê có trong Khung Chương trình giáo dục phổ thông môn toán 2018 với nội dung thống kê môn Toán của Khan Academy.

2.1. Xem xét nội dung kiến thức thống kê từ lớp 6 đến lớp 9

2.1.1. Lớp 6 giảng dạy những kiến thức thống kê: thu thập, tổ chức, biểu diễn, phân tích và xử lý dữ liệu; bảng số liệu, biểu đồ tranh, biểu đồ cột, biểu đồ cột kép.

Về mặt nội dung, có một số khái niệm như "đối tượng thống kê", "tiêu chí thống kê" cần tranh luận.

Ví dụ trong sách Toán Cánh diều tập 2:

1. Bảng số liệu

2 Bảng số liệu đánh giá chất lượng sản phẩm bằng điểm số 1, 2, 3, 4, 5 của một lô hàng gồm 20 sản phẩm như sau:

Điểm	1	2	3	4	5
Số sản phẩm	0	0	3	5	12

Quan sát bảng số liệu trên, ta thấy:

• Đối tượng thống kê là các điểm số: 1, 2, 3, 4, 5.

Các đối tượng này lần lượt được biểu diễn ở dòng đầu tiên.

• Tiêu chí thống kê là số sản phẩm ứng với mỗi loại điểm.

• Ứng với mỗi đối tượng thống kê có một số liệu thống kê theo tiêu chí, lần lượt được biểu diễn ở dòng thứ hai (theo cột tương ứng).

Trong khoa học máy tính, một đối tượng (tiếng Anh: object) có thể là một biến, một cấu trúc dữ liệu, một hàm, hay một phương thức, và như vậy, là một vị trí trong bộ nhớ có giá trị và có thể được tham chiếu bởi một định danh[1].

Trong quản lý cơ sở dữ liệu quan hệ, một đối tượng có thể là một bảng hoặc cột, hay là sự liên kết giữa dữ liệu và một thực thể cơ sở dữ liệu (ví dụ như quan hệ giữa tuổi của người với một người cụ thể)[2].

Trong ngữ cảnh thống kê, "đối tượng thống kê" là các hiện tượng, sự kiện, hoặc vận hành mà chúng ta muốn nghiên cứu và thu thập dữ liệu để phân tích. Đối tượng thống kê có thể là bất kỳ thứ gì từ một nhóm người, một quy trình sản xuất, đến một xu hướng kinh tế hoặc một hiện tượng xã

[1] https://vi.wikipedia.org/wiki/

[2] *Oppel, Andy (2005). SQL Demystified. McGraw Hill. tr. 7. ISBN 0-07-226224-9.*

hội[3]. Nói cách khác, đối tượng thường được hiểu là nhóm hoặc tập hợp mà chúng ta muốn nghiên cứu, thu thập dữ liệu và rút ra các kết luận thống kê. Như vậy, trong ví dụ được nêu trên "đối tượng thống kê" nên được coi như cả bảng hay mỗi cột số liệu bao gồm giá trị nhận được về điểm và số sản phẩm, chứ không nên phân tách bảng số liệu theo dòng "đối tượng thống kê" và "tiêu chí thống kê" như trong sách.

Tiêu chí là một tiêu chuẩn hoặc quy tắc được sử dụng để đánh giá, so sánh hoặc lựa chọn giữa các tùy chọn khác nhau. Nó có thể áp dụng trong nhiều lĩnh vực, từ kinh doanh đến khoa học và xã hội. Trong ngữ cảnh thống kê, "tiêu chí" thường được hiểu là các yếu tố hoặc đặc điểm cụ thể được sử dụng để đánh giá, so sánh, hoặc đưa ra quyết định về một tập hợp dữ liệu hoặc kết quả nghiên cứu. Các tiêu chí này có thể bao gồm các chỉ số như trung bình, trung vị, phương sai, độ lệch chuẩn, và các giá trị p trong kiểm định giả thuyết[4]. Nên chăng sử dụng cụm từ "tiêu thức thống kê" - chỉ đặc điểm của đơn vị - thay cho cụm từ "tiêu chí thống kê" trong trường hợp này.

Trong tài liệu "Thuật ngữ và khái niệm thống kê" của Cục Thống kê Úc[5] có minh họa bảng dữ liệu như sau:

Ví dụ về tập dữ liệu

	age (years)	sex	income ($)	
Person 1 (John Smith)	18	m	50000	Data Items
Person 2 (Joe Bloggs)	19	m	40000	Data Unit - Person 2
Person 3 (Sally Jones)	20	f	55000	
Person 4 (Linda Lee)	22	f	50000	Numeric observation of the data item 'income'
Person 5 (Harry James)	19	m	35000	Non-numeric (categorical) observation of the data item 'sex'

Dữ liệu (data) là các đo lường hoặc quan sát được thu thập như một nguồn thông tin. Có nhiều loại dữ liệu khác nhau và các cách khác nhau để biểu diễn dữ liệu.

Đơn vị dữ liệu (data unit) là một thực thể (entity) chẳng hạn như một người hoặc doanh nghiệp trong tổng thể (population) đang được nghiên cứu, về dữ liệu nào được thu thập. Đơn vị dữ liệu còn được gọi là bản ghi hoặc bản ghi đơn vị.

Mục dữ liệu (data item) là một đặc điểm (hoặc thuộc tính) của một đơn vị dữ liệu được đo lường hoặc đếm, chẳng hạn như chiều cao, quốc tịch hoặc thu nhập. Một mục dữ liệu cũng được gọi là một biến vì đặc tính có thể khác nhau giữa các đơn vị dữ liệu và có thể thay đổi theo thời gian.

Quan sát (observation) là sự xuất hiện của một mục dữ liệu cụ thể được ghi lại về một đơn vị dữ liệu. Nó cũng có thể được gọi là dữ liệu dạng số ít (datum). Một quan sát có thể là số hoặc không phải số (phân loại).

Một tập dữ liệu là một tập hợp đầy đủ của tất cả các quan sát.

Tôi muốn đưa ra những khái niệm và thuật ngữ từ những nguồn đáng tin cậy để so sánh và chuẩn hóa khái niệm "đối tượng thống kê" và "tiêu chí thống kê" có trong sách giáo khoa toán lớp 6.

2.1.2. Lớp 7 giảng dạy kiến thức thống kê về thu thập, phân loại và biểu diễn dữ liệu; phân tích và xử lí dữ liệu; biểu đồ đoạn thẳng; biểu đồ hình quạt tròn.

2.1.2.1. Phân loại dữ liệu, trong sách viết: "Trong các dữ liệu thống kê thu thập được, có những dữ liệu thống kê là số (số liệu) nhưng cũng có những dữ liệu thống kê không phải là số".

Ví dụ trong sách:

I. THU THẬP VÀ PHÂN LOẠI DỮ LIỆU

1 Đọc kĩ các nội dung sau:

Lớp trưởng lớp 7D thu thập thông tin về Tổ I được những dữ liệu thống kê sau:

– Tổ I gồm mười bạn, đó là: An, Bích, Châu, Chung, Dung, Dương, Quỳnh, Sơn, Thuỷ, Việt.

– Số đo chiều cao (theo đơn vị xăng-ti-mét) của mười bạn đó lần lượt là: 153, 150, 154, 151, 152, 152, 154, 156, 155, 154.

Nhận xét: Trong các dữ liệu thống kê thu thập được, có những dữ liệu thống kê là số (số liệu) nhưng cũng có những dữ liệu thống kê không phải là số.

[3] *Đối tượng thống kê và tiêu chí thống kê là gì (accgroup.vn)*

[4] *Tiêu chí thống kê là gì - Thống kê là gì - Đối tượng thống kê là gì (tìm hiểu các thông tin và vấn đề liên quan) - BYTUONG*

[5] *Statistical terms and concepts glossary | Australian Bureau of Statistics (abs.gov.au)*

Theo tôi, trong phân loại dữ liệu, nên sử dụng thuật ngữ "định lượng" và "định tính" thay vì thuật ngữ "dữ liệu số (số liệu)" và "dữ liệu không phải số". Trong "Thuật ngữ và khái niệm thống kê" của Cục Thống kê Úc dữ liệu được phân loại thành dữ liệu định lượng và dữ liệu định tính.

Định lượng = Số lượng. Dữ liệu định lượng là:

- đo lường các giá trị hoặc số lượng và được biểu thị dưới dạng số.
- dữ liệu về các biến số (ví dụ: bao nhiêu, hoặc bao nhiêu lần).

Định tính = Chất lượng. Dữ liệu định tính là:

- đo lường 'phân loại' và có thể được biểu thị bằng tên, ký hiệu hoặc mã số.
- dữ liệu về các biến phân loại (ví dụ: loại nào).

Lên lớp 8 thuật ngữ "định lượng" và "định tính" mới được sử dụng.

Trong các dữ liệu thống kê thu thập được:
– Có những dữ liệu thống kê là số (số liệu), những dữ liệu này còn gọi là *dữ liệu định lượng*;
– Có những dữ liệu thống kê không phải là số, những dữ liệu này còn gọi là *dữ liệu định tính*.

Nếu ta coi dữ liệu số là định lượng, dữ liệu không phải số là định tính thì sự phân loại đó chưa hẳn đúng, ví dụ, xếp hạng học lực học sinh theo thứ tự từ 1 đến 10 là dữ liệu dạng số, nhưng nó không phải là dữ liệu định lượng, mà là dữ liệu định tính biểu thị thứ bậc. Cần chuẩn hóa việc phân loại dữ liệu ngay từ ban đầu.

2.1.2.2. Tính hợp lí của kết luận thống kê

Ví dụ trong sách:

II. TÍNH HỢP LÍ CỦA KẾT LUẬN THỐNG KÊ

2 Đọc kĩ các nội dung sau:

Quá trình phân tích và xử lí dữ liệu giúp chúng ta có thể nhận biết được: tính hợp lí của dữ liệu thống kê, tính hợp lí của kết luận thống kê và ta cũng có thể bác bỏ kết luận đã nêu ra. Thông thường, để làm được điều đó ta dựa trên những tiêu chí đơn giản hoặc dựa trên tính toán và suy luận toán học.

Ở đây hình như có sự hiểu đồng nhất tính hợp lí của dữ liệu với tính hợp lí của kết luận thống kê. Hơn nữa hiểu "những tiêu chí đơn giản dựa trên tính toán và suy luận toán học" là như thế nào? Sách giáo khoa chưa đưa ra những dẫn giải cụ thể về "những tiêu chí toán học đơn giản" mà chỉ có minh họa qua một số ví dụ.

Tính hợp lý (validity) của dữ liệu đề cập đến tính đúng đắn (đo lường chính xác) và tính liên quan của dữ liệu (những gì nó dự định đo lường). Dữ liệu hợp lí phải phù hợp và đại diện cho các số liệu được thu thập hoặc thế giới thực mà nó mô tả. Nói cách khác, dữ liệu hợp lí thể hiện chính xác cấu trúc, thuộc tính hoặc hiện tượng trong thế giới thực mà nó được thiết kế để nắm bắt[6].

Có một số phương pháp xác định tính hợp lí của dữ liệu, bao gồm[7]:

- Xác thực: Kiểm tra xem thông tin đã nhập có tuân theo các quy tắc cụ thể hay không, chẳng hạn như đảm bảo số điện thoại có đúng số chữ số; đúng định dạng - dữ liệu phải tuân theo định dạng đã định trước. Ví dụ, họ và tên phải là chữ, số tuổi phải là số, email phải có kí hiệu @, v.v…
- Đánh giá: Kiểm tra kỹ thông tin, chẳng hạn xem qua ghi chú để tìm lỗi; kiểm tra phạm vi dự kiến - dữ liệu phải nằm trong khoảng giá trị dự kiến. Ví dụ, số người không thể âm, cân nặng của người Việt Nam không thể vượt quá 200kg.
- So sánh: Tham khảo chéo thông tin với các nguồn khác, như so sánh ghi chú với sách giáo khoa để đảm bảo các sự kiện là chính xác; xét tính đại diện - dữ liệu phải đại diện cho vấn đề cần thống kê. Điều này đảm bảo rằng kết quả từ dữ liệu có ý nghĩa và chính xác.

"Những tiêu chí toán học đơn giản" để nhận biết được tính hợp lí của dữ liệu như sách đã dẫn

[6] *Reliability vs. Validity in Research | Difference, Types and Examples (scribbr.com)*

[7] *Tính toàn vẹn dữ liệu so với tính hợp lệ của dữ liệu: Sự khác biệt chính với sự tương tự của sở thú (montecarlodata.com)*

chưa được dẫn giải theo một số phương pháp nêu trên. Trong sách việc thực hiện xem xét tính hợp lí của dữ liệu nặng về thực hiện các phép tính.

Kết luận thống kê là phần quan trọng trong quá trình phân tích dữ liệu. Nó giúp chúng ta rút ra những thông tin quan trọng từ dữ liệu và đưa ra những kết luận có ý nghĩa[8]. Kết luận thống kê không chỉ là việc đưa ra con số, mà còn là việc hiểu và diễn giải kết quả một cách hợp lý. Như vậy không thể chỉ "dựa trên những tiêu chí đơn giản hoặc dựa trên tính toán và suy luận toán học" để nhận biết được tính hợp lý của kết luận thống kê.

Kết luận có ý nghĩa hay không thường dựa trên kiểm định giả thuyết. Kiểm định giả thuyết là một trong những phương pháp của Thống kê suy luận - đưa ra kết luận hoặc đưa ra dự đoán về tổng thể dựa trên dữ liệu mẫu, sử dụng lý thuyết xác suất để ước lượng và kiểm định giả thuyết.

Trong khi nội dung kiến thức thống kê được giảng dạy trong chương trình giáo dục phổ thông hiện nay được giới hạn trong khuôn khổ Thống kê mô tả - sử dụng thống kê để tổng hợp và lập bảng, biểu đồ cho một tập dữ liệu, làm cho các đặc trưng của nó (trung bình, phương sai, phân bố…) biểu lộ gọn và rõ hơn. Việc xem xét tính hợp lí của kết luận thống kê đòi hỏi kiến thức vượt quá phạm vi môn học thống kê đối với học sinh phổ thông.

2.1.3. Lớp 8 giảng dạy kiến thức thống kê sau: thu thập và phân loại dữ liệu; mô tả và biểu diễn dữ liệu trên các bảng, biểu đồ; phân tích và xử lí dữ liệu.

2.1.3.1. Thu thập dữ liệu

Ví dụ trong sách:

I. THU THẬP DỮ LIỆU

1 Các bạn học sinh lớp 8A muốn thu thập thông tin về số lượng huy chương đạt được của Đoàn thể thao Việt Nam tại SEA Games 30. Theo em, các bạn lớp 8A có thể thu thập những thông tin đó bằng cách nào?

Có nhiều cách để thu thập dữ liệu, chẳng hạn: quan sát, lập phiếu điều tra (phiếu hỏi), tiến hành phỏng vấn, ... hoặc thu thập từ những nguồn có sẵn như sách, báo, trang web, các phương tiện thông tin đại chúng, ...

Thu thập dữ liệu là quá trình thu thập thông tin từ các nguồn có liên quan để tìm ra giải pháp cho yêu cầu thống kê nhất định[9]. Dữ liệu có thể được quan sát, đo lường hoặc thu thập bằng cách đặt câu hỏi phỏng vấn trong khảo sát/ điều tra hoặc tổng điều tra[10].

Dữ liệu có thể được phân thành hai loại: dữ liệu sơ cấp và dữ liệu thứ cấp. Dữ liệu sơ cấp là dữ liệu được thu thập từ các nguồn sơ cấp lần đầu tiên thông qua quan sát, đo lường hay khảo sát, điều tra. Còn dữ liệu thứ cấp là dữ liệu đã tồn tại được người khác thu thập trước đó cho các mục đích khác. Dữ liệu thứ cấp có thể được thu thập thông qua các nguồn được xuất bản và chưa được công bố khác nhau (có sẵn).

Trong sách giáo khoa viết "Có nhiều cách để thu thập dữ liệu …" mà chưa phân biệt rõ loại dữ liệu nào (sơ cấp hay thứ cấp) được thu thập theo những cách nào!

Yêu cầu trong Khung Chương trình giáo dục phổ thông môn toán lớp 8 là "Thực hiện và lí giải được việc thu thập, phân loại dữ liệu theo các tiêu chí cho trước từ nhiều nguồn khác nhau: văn bản; bảng biểu; kiến thức trong các lĩnh vực giáo dục khác (Địa lí, Lịch sử, Giáo dục môi trường, Giáo dục tài chính,...); phỏng vấn, truyền thông, Internet; thực tiễn (môi trường, tài chính, y tế, giá cả thị trường,...)". Yêu cầu này không tách bạch rõ giữa nguồn dữ liệu với cách thức (phương pháp) thu thập dữ liệu.

Nguồn dữ liệu (data source) thường được sử dụng như một từ đồng nghĩa với thuật ngữ "nhà

[8] *Ý nghĩa thống kê (Statistical Significance) là gì? Đặc điểm (vietnambiz.vn)*

[9] *Thu thập dữ liệu và phương pháp của nó - GeeksforGeeks*

[10] *OECD Glossary of Statistical Terms | READ online (oecd-ilibrary.org)*

cung cấp dữ liệu" - nơi có sẵn dữ liệu hoặc siêu dữ liệu. Tuy nhiên, trong ngữ cảnh Trao đổi dữ liệu thống kê và siêu dữ liệu (SDMX), thuật ngữ này đề cập đến tổ chức hoặc cá nhân nơi thu thập số liệu thống kê[11].

Còn về phương pháp thu thập dữ liệu, có hai phương pháp thu thập dữ liệu khác nhau: thu thập dữ liệu sơ cấp và thu thập dữ liệu thứ cấp như đã nói ở trên[12].

Nội dung kiến thức về Thu thập dữ liệu trong sách giáo khoa đã không tách biệt rõ ràng nguồn dữ liệu với cách thức thu thập dữ liệu.

2.1.3.2. Tính hợp lí của dữ liệu

Ví dụ trong sách:

Nhận xét: Để đánh giá tính hợp lí của dữ liệu, ta cần đưa ra các tiêu chí đánh giá, chẳng hạn như dữ liệu phải:

– Đúng định dạng;

– Nằm trong phạm vi dự kiến;

– Phải có tính đại diện đối với vấn đề cần thống kê.

Sau đó lại dẫn giải đến các tiêu chí toán học:

Để đánh giá tính hợp lí của dữ liệu, ta có thể dựa vào mối liên hệ toán học đơn giản giữa các số liệu.

Chẳng hạn, để đảm bảo tính hợp lí, dữ liệu cần phải đáp ứng đúng các tiêu chí toán học đơn giản như:

– Tổng tất cả các số liệu thành phần phải bằng số liệu của toàn thể;

– Số lượng của bộ phận phải nhỏ hơn số lượng của toàn thể.

Nội dung kiến thức trong sách không có chỉ dẫn liên quan giữa tiêu chí đánh giá với tiêu chí toán học nêu trên. Tại Mục 2.1.2.2 ở trên đã chỉ ra "những tiêu chí toán học đơn giản" để thực hiện đánh giá tính hợp lí của dữ liệu như sách đã dẫn chưa được dẫn giải theo một số phương pháp (xác thực, đánh giá, so sánh) để xác định tính hợp lí của dữ liệu mà còn nặng về công thức toán học thuần túy.

2.1.3.3. Lựa chọn và biểu diễn dữ liệu vào bảng, biểu đồ thích hợp

Ví dụ trong sách:

Nhận xét

– Để biểu diễn dữ liệu thống kê, ta cần lựa chọn bảng, biểu đồ thích hợp.

– Để có thể hoàn thiện được biểu đồ thống kê (hoặc bảng thống kê) đã lựa chọn, ta cần biểu diễn được dữ liệu vào biểu đồ (hoặc bảng) đó. Muốn vậy, ta cần biết cách xác định mỗi yếu tố của biểu đồ (hoặc bảng) thống kê đó.

[11] *OECD Glossary of Statistical Terms | READ online (oecd-ilibrary.org)*

[12] *Thu thập dữ liệu và phương pháp của nó - GeeksforGeeks*

Ví dụ 5 Số lượng học sinh ở hai lớp 8A và 8B đăng kí tham gia: câu lạc bộ (CLB) bóng đá lần lượt là 26 và 22 (học sinh); CLB bóng rổ lần lượt là 14 và 18 (học sinh).

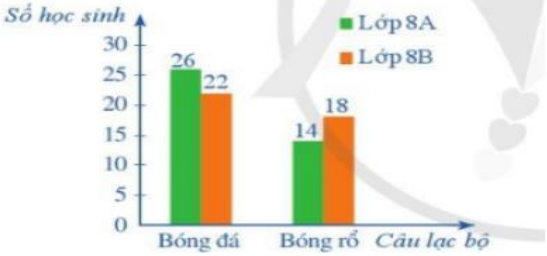
a) Hãy lựa chọn biểu đồ thích hợp để biểu diễn dữ liệu trên.

b) Hãy hoàn thiện biểu đồ ở *Hình 8* để nhận được biểu đồ cột kép thống kê số lượng học sinh ở hai lớp 8A và 8B đăng kí tham gia hai CLB trên.

Giải

a) Ta lựa chọn biểu đồ cột kép để biểu diễn dữ liệu trên.

b) Sau khi hoàn thiện biểu đồ ở *Hình 8*, ta nhận được biểu đồ cột kép ở *Hình 9* biểu diễn số lượng học sinh ở hai lớp 8A và 8B đăng kí tham gia CLB bóng đá và CLB bóng rổ.



Trong các ví dụ đã dẫn chưa thấy giải thích tại sao lại lựa chọn biểu diễn dữ liệu vào bảng hay biểu đồ nào là thích hợp!

2.1.3.4. Biểu diễn một tập dữ liệu theo những cách khác nhau

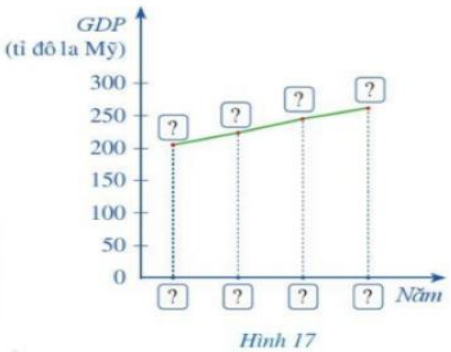
Ví dụ trong sách:

Ví dụ 9 Như đã nêu trong *Ví dụ 2*, tổng sản phẩm trong nước (GDP) theo giá hiện hành của Việt Nam trong các năm 2016, 2017, 2018, 2019 lần lượt là: 205,3; 223,7; 245,2; 261,9 (đơn vị: tỉ đô la Mỹ).

a) Lập bảng số liệu thống kê GDP của Việt Nam trong các năm trên theo mẫu sau:

Năm	2016	2017	2018	2019
GDP (tỉ đô la Mỹ)	?	?	?	?

b) Hãy hoàn thiện biểu đồ ở *Hình 17* để nhận được biểu đồ đoạn thẳng biểu diễn GDP của Việt Nam trong các năm trên.



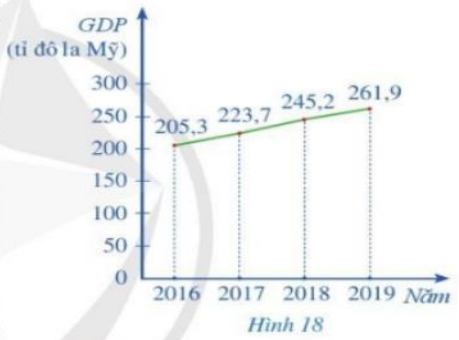
Hình 17

Giải

a) Ta có bảng số liệu thống kê GDP của Việt Nam trong các năm đã nêu là:

Năm	2016	2017	2018	2019
GDP (tỉ đô la Mỹ)	205,3	223,7	245,2	261,9

b) Sau khi hoàn thiện biểu đồ ở *Hình 17*, ta nhận được biểu đồ đoạn thẳng ở *Hình 18* biểu diễn GDP của Việt Nam trong các năm trên.



Hình 18

Nhận xét: Đối với một tập dữ liệu, ta có thể:

– Biểu diễn tập dữ liệu đó theo những cách khác nhau vào bảng, biểu đồ thích hợp;

– Chuyển tập dữ liệu đó từ dạng biểu diễn này sang dạng biểu diễn khác.

Trong khi Khung Chương trình giáo dục phổ thông môn toán lớp 8 yêu cầu:

– So sánh được các dạng biểu diễn khác nhau cho một tập dữ liệu.

– Mô tả được cách chuyển dữ liệu từ dạng biểu diễn này sang dạng biểu diễn khác.

Như vậy giữa yêu cầu kiến thức cần đạt được của Chương trình giáo dục phổ thông môn toán lớp 8 với nội dung kiến thức giảng dạy có trong sách giáo khoa còn có một khoảng cách nhất định.

2.1.3.5. Phát hiện vấn đề dựa trên phân tích và xử lí dữ liệu thu được ở dạng bảng, biểu đồ

Ví dụ trong sách:

Để phát hiện vấn đề (hoặc quy luật đơn giản) dựa trên phân tích và xử lí số liệu thu được, ta cần:
– Nhận biết được mối liên hệ toán học đơn giản giữa các số liệu đã được biểu diễn;
– Thực hiện được tính toán và suy luận toán học.

Ví dụ 3 Biểu đồ hình quạt tròn ở *Hình 30* biểu diễn kết quả thống kê (tính theo tỉ số phần trăm) kế hoạch chi tiêu hàng tháng của gia đình bác Hạnh.

a) Khoản chi tiêu nào của gia đình bác Hạnh là lớn nhất?

b) Số tiền chi tiêu hàng tháng của gia đình bác Hạnh dành cho ăn uống gấp bao nhiêu lần số tiền dành cho tiết kiệm?

c) Tính số tiền gia đình bác Hạnh tiết kiệm hàng tháng theo kế hoạch, biết tổng thu nhập hàng tháng của gia đình bác Hạnh là 25 triệu đồng.

Hình 30

Giải

a) Khoản chi tiêu hàng tháng dành cho ăn uống của gia đình bác Hạnh là lớn nhất.

b) Do 35 : 20 = 1,75 nên số tiền chi tiêu hàng tháng của gia đình bác Hạnh dành cho ăn uống gấp 1,75 lần số tiền dành cho tiết kiệm.

c) Số tiền gia đình bác Hạnh tiết kiệm hàng tháng theo kế hoạch là: 25 . 20% = 5 (triệu đồng).

Phát hiện vấn đề hoặc quy luật từ dữ liệu đòi hỏi sự phân tích kỹ lưỡng và hiểu biết về ngữ cảnh. Trong ví dụ 3 trên, khoản chi tiêu cho ăn uống là lớn nhất thì cần tập trung tìm hiểu nguyên nhân để giải quyết vấn đề chi tiêu lớn này, chứ không phải tiếp tục tính toán so sánh số tiền chi tiêu với số tiền tiết kiệm!

Một ví dụ cụ thể: Biểu đồ xu hướng thị trường chứng khoán

Giả sử bạn đang nghiên cứu biểu đồ xu hướng giá cổ phiếu của một công ty trong một khoảng thời gian dài. Dựa vào biểu đồ, bạn thấy rằng giá cổ phiếu liên tục giảm trong 6 tháng qua. Điều này có thể là dấu hiệu của một vấn đề tiềm ẩn:

(i) Suy thoái kinh tế: Nếu giá cổ phiếu giảm liên tục, có thể thị trường đang trải qua suy thoái kinh tế. Điều này có thể ảnh hưởng đến doanh nghiệp và lợi nhuận của họ.

(ii) Vấn đề nội bộ: Có thể công ty đang gặp vấn đề nội bộ, chẳng hạn như thay đổi lãnh đạo, thất thoát tài sản, hoặc không thể duy trì hiệu suất kinh doanh.

(iii) Thay đổi trong ngành: Nếu toàn ngành đang gặp khó khăn, giá cổ phiếu của công ty cũng có thể bị ảnh hưởng.

(iv) Thay đổi chính sách: Có thể có thay đổi trong chính sách kinh tế, thuế, hoặc quy định liên quan đến công ty.

Bằng cách xem xét các số liệu thu được ở dạng bảng hay biểu đồ có thể so sánh, tìm ra xu hướng hay biến động tăng/ giảm, đó có thể là một vấn đề hoặc quy luật để tìm hiểu nguyên nhân hoặc đưa ra được kết luận về dữ liệu đó. Chứ không phải "nhận biết được mối liên hệ toán học … để phát hiện vấn đề (hoặc quy luật đơn giản)" như sách đã dẫn.

2.1.3.6. Giải quyết những vấn đề đơn giản dựa trên phân tích và xử lí dữ liệu thu được ở dạng bảng, biểu đồ

Câu hỏi đặt ra "Giải quyết những vấn đề được phát hiện (dựa trên phân tích và xử lí số liệu thu được) là gì?"

Câu trả lời là: "Giải quyết vấn đề dựa trên phân tích và xử lý số liệu thu được là quá trình tìm ra giải pháp cho một tình huống hoặc vấn đề cụ thể".

Ví dụ trong sách:

Để giải quyết vấn đề đã được phát hiện (dựa trên phân tích và xử lí số liệu thu được), ta cần thực hiện những tính toán và suy luận trên cơ sở mối liên hệ toán học giữa các số liệu đó.

Ví dụ 4 Biểu đồ cột ở *Hình 31* thống kê sản lượng sản xuất than ở tỉnh Quảng Ninh trong các năm 2017, 2018, 2019, 2020. Căn cứ vào thống kê trên, một bài báo đã nêu ra nhận định: "Tổng sản lượng sản xuất than ở tỉnh Quảng Ninh trong các năm 2017, 2018, 2019, 2020 đã đạt xấp xỉ 164 triệu tấn và so với năm 2017, sản lượng sản xuất than ở tỉnh Quảng Ninh trong năm 2020 đã tăng lên xấp xỉ 34%". Em hãy cho biết nhận định trên của bài báo có chính xác không?

Giải

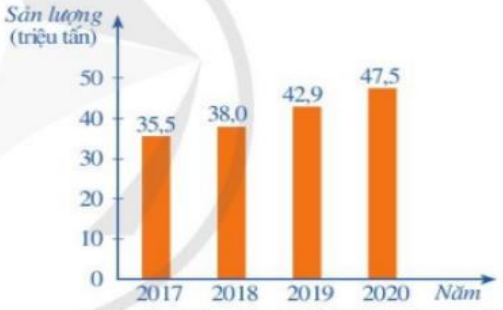
Tổng sản lượng sản xuất than ở tỉnh Quảng Ninh trong các năm 2017, 2018, 2019, 2020 là:

$$35,5 + 38 + 42,9 + 47,5 \approx 164 \text{ (triệu tấn)}.$$

Ở tỉnh Quảng Ninh, tỉ số phần trăm của sản lượng sản xuất than trong năm 2020 và sản lượng sản xuất than trong năm 2017 là:

$$\frac{47,5 \cdot 100}{35,5}\% \approx 134\%.$$

Vậy nhận định trên của bài báo là chính xác.



(Nguồn: Báo cáo của Bộ Công thương từ năm 2017 đến năm 2020)

Hình 31

Trong ví dụ Biểu đồ xu hướng thị trường chứng khoán nêu trên, từ việc phát hiện vấn đề qua phân tích dữ liệu, chúng ta có thể giải quyết vấn đề bằng cách thực hiện các giải pháp cho vấn đề cụ thể:

(i) Suy thoái kinh tế: Dựa trên phân tích dữ liệu, xác định chiến lược giao dịch phù hợp. Cân nhắc các điểm vào/thoát lệnh và quản lý rủi ro.

(ii) Vấn đề nội bộ: Điều tra nguyên nhân gây ra vấn đề. Có thể là do thay đổi trong chính sách, sự kiện kinh tế, hoặc yếu tố khác ảnh hưởng đến thị trường chứng khoán.

(iii) Thay đổi trong ngành: Theo dõi hiệu suất của chiến lược và điều chỉnh khi cần thiết. Điều này đòi hỏi việc liên tục theo dõi biểu đồ và cập nhật thông tin.

(iv) Thay đổi chính sách: Theo dõi thay đổi trong chính sách kinh tế, thuế, hoặc quy định liên quan đến công ty. Điều này đòi hỏi việc liên tục theo dõi biểu đồ và cập nhật thông tin.

Giải quyết những vấn đề được phát hiện qua phân tích số liệu thu được ở dạng bảng hay biểu đồ là đưa ra các giả thuyết về nguyên nhân và giải pháp cho vấn đề đó. Đây là kiến thức khó đối với học sinh phổ thông.

Nhớ rằng, việc phân tích số liệu từ các biểu đồ và bảng thống kê đòi hỏi sự kỹ lưỡng và kiên nhẫn để hiểu rõ hơn về các vấn đề và quy luật trong dữ liệu, từ đó đưa ra các giả thuyết về nguyên nhân và tìm ra giải pháp cho một tình huống hoặc vấn đề cụ thể.

Nội dung kiến thức trong sách chưa đưa ra được những ví dụ sinh động về phát hiện và giải quyết được vấn đề hoặc quy luật đơn giản dựa trên phân tích các số liệu thu được ở dạng bảng hay biểu đồ theo cách hiểu chuẩn tắc được minh họa bằng ví dụ cụ thể về Biểu đồ xu hướng thị trường chứng khoán mà tôi đã đưa ra.

2.1.4. Lớp 9 giảng dạy kiến thức thống kê liên quan đến mô tả và biểu diễn dữ liệu trên các bảng, biểu đồ; tần số, tần số tương đối; tần số ghép nhóm, tần số tương đối ghép nhóm.

2.1.4.1. Biểu diễn dữ liệu trên bảng thống kê, biểu đồ tranh, biểu đồ cột, biểu đồ đoạn thẳng, biểu đồ hình quạt tròn

Ví dụ trong sách:

Thuật ngữ "đối tượng thống kê" và "tiêu chí thống kê" là những thuật ngữ đã được tranh luận tại Mục 1 Phần I trên. Theo tôi một bảng thống kê nên được trình bày như ví dụ về tập dữ liệu.

Để biểu diễn dữ liệu trên bảng thống kê, ta có thể làm như sau:

Bước 1. Các đối tượng thống kê lần lượt được biểu diễn ở cột đầu tiên, trong khi các tiêu chí thống kê lần lượt được biểu diễn ở dòng đầu tiên hoặc ngược lại

Bước 2. Các số liệu thống kê theo tiêu chí của mỗi đối tượng thống kê lần lượt được biểu diễn ở dòng (hoặc cột) tương ứng.

Ví dụ về tập dữ liệu

	age (years)	sex	income ($)
Person 1 (John Smith)	18	m	50000
Person 2 (Joe Bloggs)	18	m	40000
Person 3 (Sally Jones)	20	f	55000
Person 4 (Linda Lee)	22	f	50000
Person 5 (Harry James)	19	m	35000

Data Items

Data Unit - Person 2.

Numeric observation of the data item 'income'

Non-numeric (categorical) observation of the data item 'sex'

Thông thường, dữ liệu được sắp xếp theo dạng bảng, trong đó mỗi cột của bảng đại diện cho một biến hay một mục (loại số liệu) cụ thể và mỗi hàng tương ứng với một bản ghi hay một thực thể đối tượng nhất định của tập dữ liệu. Dữ liệu (giá trị quan sát) được ghi vào ô giao giữa các cột và các hàng tương ứng như hình minh họa.

Tập dữ liệu có thể được biểu diễn bằng một số dạng biểu đồ phổ biến sau:

(i) Biểu đồ tranh: Biểu thị các giá trị định lượng bằng biểu tượng hay hình ảnh một cách trực quan.

Thuật ngữ "đối tượng thống kê" trong nội dung kiến thức này, nên được thay thế bởi cụm từ "thực thể" hay "tên đối tượng".

Để biểu diễn dữ liệu trên biểu đồ tranh, ta có thể làm như sau:

Bước 1. Các đối tượng thống kê được biểu diễn ở cột đầu tiên của bảng thống kê

Bước 2. Chọn biểu tượng để biểu diễn số liệu thống kê. Các biểu tượng đó được trình bày ở dòng cuối cùng trong bảng thống kê

Bước 3. Số liệu thống kê theo tiêu chí của mỗi đối tượng thống kê được biểu diễn bằng các biểu tượng ở dòng tương ứng trong bảng thống kê.

(ii) Biểu đồ cột: Sử dụng cho việc so sánh trực quan giữa các nhóm/ mục dữ liệu hoặc theo thời gian.

Trong nội dung kiến thức, thuật ngữ "đối tượng thống kê" nên được thay thế bởi cụm từ "thực thể" hay "tên đối tượng"; "tiêu chí" nên được thay bởi cụm từ "tiêu thức" hay "mục dữ liệu" hoặc "biến".

Để biểu diễn dữ liệu trên biểu đồ cột, ta có thể làm như sau:

Bước 1. Vẽ hai trục vuông góc với nhau

- Trên trục nằm ngang: biểu diễn các đối tượng thống kê
- Trên trục thẳng đứng: xác định độ dài đơn vị để biểu diễn số liệu thống kê và cần chọn độ dài đơn vị thích hợp với số liệu

Bước 2. Tại vị trí các đối tượng thống kê trên trục nằm ngang, vẽ những cột hình chữ nhật: cách đều nhau; có cùng chiều rộng; có chiều cao thể hiện số liệu thống kê theo tiêu chí của mỗi đối tượng thống kê

Bước 3. Hoàn thiện biểu đồ: ghi tên các trục và ghi số liệu tương ứng trên mỗi cột (nếu cần).

(iii) Biểu đồ đoạn thẳng: Sử dụng khi biểu diễn sự thay đổi/ xu hướng của chuỗi dữ liệu theo thời gian.

Trong nội dung kiến thức, thuật ngữ "đối tượng thống kê" nên được thay thế bởi cụm từ "thực thể" hay "tên đối tượng"; "tiêu chí" nên được thay bởi cụm từ "tiêu thức" hay "mục dữ liệu" hoặc "biến".

Để biểu diễn dữ liệu trên biểu đồ đoạn thẳng, ta có thể làm như sau:

Bước 1. Vẽ hai trục vuông góc với nhau tại điểm O

- Trên trục nằm ngang: mỗi đối tượng thống kê được đánh dấu bằng một điểm và các điểm này thường được vẽ cách đều nhau
- Trên trục thẳng đứng: xác định độ dài đơn vị để biểu diễn số liệu thống kê và cần chọn độ dài đơn vị thích hợp với số liệu, đánh dấu điểm theo tiêu chí của đối tượng thống kê tương ứng

Bước 2. Với mỗi đối tượng thống kê, ta tiếp tục:

- Xác định điểm A đánh dấu số liệu thống kê trên trục thẳng đứng của đối tượng thống kê đó
- Kẻ bằng nét đứt một đoạn thẳng có độ dài bằng *OA*, vuông góc với trục nằm ngang và đi qua điểm đánh dấu đối tượng thống kê đó trên trục nằm ngang. Đầu mút trên của đoạn thẳng đó là điểm mốc của đối tượng thống kê

Bước 3. Vẽ đường gấp khúc gồm các đoạn thẳng nối liền liên tiếp các điểm mốc

Bước 4. Hoàn thiện biểu đồ: ghi tên các trục và ghi số liệu tương ứng trên mỗi điểm mốc (nếu cần).

(iv) Biểu đồ hình quạt tròn: Sử dụng khi muốn so sánh tỷ lệ phần trăm của các phần/ mục dữ liệu so với tổng thể.

Trong nội dung kiến thức, thuật ngữ "đối tượng thống kê" nên được thay thế bởi từ "nhãn (label)" hay "tên đối tượng".

Chúng ta nên sử dụng những từ ngữ chuẩn có trong tài liệu "Thuật ngữ và khái niệm thống kê" của các Cơ quan Thống kê như Thống kê Canada hay Úc, hoặc trong "Bảng chú giải thuật ngữ thống kê" của Tổ chức Hợp tác và Phát triển Kinh tế (OECD).

Để vẽ biểu đồ hình quạt tròn biểu diễn các số liệu thống kê tính theo tỉ số phần trăm, ta có thể làm như sau:

Bước 1. Vẽ đường tròn tâm O bán kính R

Bước 2. Chuyển đổi số liệu của một đối tượng thống kê (tính theo tỉ số phần trăm) về số đo cung tương ứng với đối tượng thống kê đó (tính theo độ) dựa trên nguyên tắc sau: $x\%$ tương ứng với $x\% \,.\, 360^o$

Các số đo cung tương ứng với các đối tượng thống kê được cho ở bảng sau:

Đối tượng thống kê	1	2	...	k
Số đo cung tương ứng (đơn vị: độ)	n_1	n_2	...	n_k

Bảng 7

Chú ý: $n_1 + n_2 + \ldots + n_k = 360^o$

Bước 3. – Vẽ tia gốc OA theo phương thẳng đứng

– Căn cứ vào *Bảng 7*, sử dụng thước thẳng và thước đo độ, vẽ theo chiều quay của kim đồng hồ các cung $AA_1, A_1A_2, \ldots, A_{k-2}A_{k-1}$ lần lượt có số đo là $n_1, n_2, \ldots, n_{k-1}$. Khi đó cung $A_{k-1}A$ có số đo là:

$$360^o - (n_1 + n_2 + \ldots + n_{k-1}) = n_k$$

Bước 4. Hoàn thiện biểu đồ: ghi tên đối tượng thống kê vào hình quạt tương ứng; ghi số liệu tương ứng trên mỗi hình quạt; các hình quạt được tô màu khác nhau (nếu cần) và xoá đi những thông tin không cần thiết trong biểu đồ.

2.1.4.2. Tần số: Tần số tương đối

Ví dụ trong sách:

Nhận xét

- Một tập gồm hữu hạn các dữ liệu thống kê được gọi là một mẫu. Số phần tử của một mẫu được gọi là kích thước mẫu (hay cỡ mẫu). Chẳng hạn, mẫu số liệu thống kê ở trên có kích thước mẫu là 60.

Khái niệm "Một tập gồm hữu hạn các dữ liệu thống kê được gọi là một mẫu" là không chuẩn tắc. Theo "Thuật ngữ và khái niệm thống kê" của Cục Thống kê Úc: "Mẫu (sample) là một tập hợp con của các đơn vị (thực thể) trong một tổng thể/ quần thể (population), được chọn để đại diện cho tất cả các đơn vị trong một tổng thể/ quần thể quan tâm. Nó là một bảng liệt kê một phần vì nó là một số lượng từ một phần của tổng thể".

Ví dụ trong sách:

Số lần xuất hiện của một giá trị trong mẫu dữ liệu thống kê được gọi là *tần số* của giá trị đó.

Định nghĩa "Tần số của một giá trị là số lần xuất hiện của một giá trị đó trong mẫu dữ liệu" là chưa chính xác. Mẫu dữ liệu có lẽ được hiểu là một tập hợp con của dữ liệu được chọn từ một tổng thể lớn hơn, đại diện cho tổng thể đó. Một tập dữ liệu có thể có một hoặc nhiều biến (mục dữ liệu) khác nhau. Khi nói đến tần số là nói đến số lần xuất hiện giá trị cụ thể của mỗi biến trong tập dữ liệu chứ không phải là tất cả các biến trong dữ liệu mẫu. Bởi vậy, định nghĩa chính xác là: "Tần số là số lần một giá trị cụ thể cho một biến (mục dữ liệu) đã được quan sát xảy ra". Cách đơn giản nhất để thể hiện tần số là về mặt tuyệt đối (tài liệu "Thuật ngữ và khái niệm thống kê" của Úc).

Trong sách giáo khoa coi tần số đương nhiên là tần số tuyệt đối và đưa ra khái niệm tần số tương đối.

Tần số tương đối f_i của giá trị x_i là tỉ số giữa tần số n_i của giá trị đó và số lượng N các dữ liệu trong mẫu dữ liệu thống kê: $f_i = \frac{n_i}{N}$.

Ta thường viết tần số tương đối dưới dạng phần trăm.

Khái niệm như vậy cũng chưa chuẩn xác khi nói về số lượng các dữ liệu trong mẫu dữ liệu. Bởi trong dữ liệu mẫu có thể có một hoặc nhiều biến. Định nghĩa chính xác là: "Tần số tương đối mô tả số lần một giá trị cụ thể cho một biến (mục dữ liệu) đã được quan sát xảy ra liên quan đến tổng số giá trị đối với biến đó. Tần số tương đối được tính bằng cách chia tần số tuyệt đối cho tổng số giá trị đối với biến đó" (tài liệu "Thuật ngữ và khái niệm thống kê").

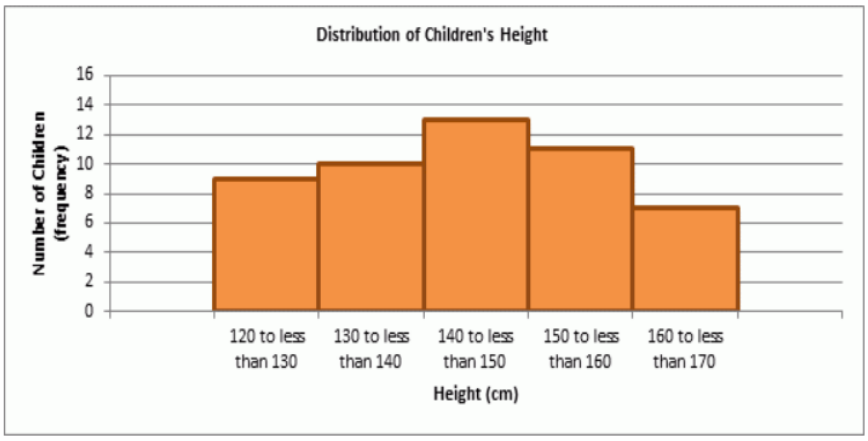
Sách giáo khoa hướng dẫn cách lập bảng tần số và tần số tương đối riêng biệt, mà chưa đề cập tới lập bảng tần số hiển thị tần số tuyệt đối và tần số tương đối cùng lúc.

Ví dụ: nếu chúng tôi đã thu thập dữ liệu về chiều cao từ một mẫu gồm 50 trẻ em, chúng tôi có thể trình bày những phát hiện của mình như bảng dưới đây.

Phân bố tần số dữ liệu có thể được hiển thị trong bảng hoặc biểu đồ. Một số phương pháp phổ biến để hiển thị phân bố tần số bao gồm bảng tần số, biểu đồ tần suất hoặc biểu đồ cột. Sách hướng dẫn vẽ biểu đồ tần số ở dạng biểu đồ cột và/hoặc biểu đồ đoạn thẳng, vẽ biểu đồ tần số tương đối ở dạng biểu đồ cột và/hoặc biểu đồ hình quạt tròn, song chưa đề cập tới biểu đồ tần suất (histogram).

Chiều cao của trẻ em

Chiều cao (cm) của trẻ em	Tần số tuyệt đối	Tần số tương đối
120 - nhỏ hơn 130	9	18%
130 – nhỏ hơn 140	10	20%
140 - dưới 150	13	26%
150 - dưới 160	11	22%
160 - dưới 170	7	14%
Tất cả	50	100%



Biểu đồ tần suất được vẽ là biểu đồ cột không có khoảng cách giữa các cột. Trục dọc thường đại diện cho tần số tuyệt đối/ tần số tương đối, trong khi trục ngang thường đại diện cho biến được đo lường. Biểu đồ cho thấy sự phân bố của tất cả các quan sát trong một tập dữ liệu định lượng. Xem ví dụ hình bên.

Tóm lại, trong nội dung kiến thức thống kê chúng ta nên sử dụng những thuật ngữ và những khái niệm chuẩn có trong các tài liệu như "Thuật ngữ và khái niệm thống kê" của các Cơ quan Thống kê Canada hay Úc, hoặc "Bảng chú giải thuật ngữ thống kê" của Tổ chức Hợp tác và Phát triển Kinh tế (OECD) để đảm bảo rằng học sinh có thể hiểu và áp dụng thống kê một cách chính xác và hiệu quả.

(Còn tiếp trong Thông tin khoa học thống kê số 4)