

Phát triển tư duy thống kê cho sinh viên đại học ngành Dược thông qua luyện tập cho sinh viên đọc hiểu, tổ chức và thu gọn, biểu diễn dữ liệu thống kê

Quách Thị Sen

Trường Đại học Dược Hà Nội
13 - 15 Lê Thánh Tông, Hoàn Kiếm, Hà Nội, Việt Nam
Email: senqtdhd@gmail.com

TÓM TẮT: Phát triển tư duy thống kê là một trong những nhiệm vụ quan trọng đối với đào tạo sinh viên đại học ngành Dược. Thông qua các quan niệm về tư duy thống kê và các mức độ tư duy thống kê mà sinh viên đại học ngành Dược cần đạt được, bài viết đề xuất một số biện pháp phát triển tư duy thống kê trong dạy học Toán - Thống kê Y Dược cho sinh viên đại học ngành Dược: Luyện tập cho sinh viên đọc hiểu dữ liệu thống kê và luyện tập cho sinh viên tổ chức, thu gọn, biểu diễn dữ liệu thống kê.

TỪ KHÓA: Tư duy thống kê; dữ liệu; luyện tập; ngành Dược; sinh viên.

→ Nhận bài 02/02/2019 → Nhận kết quả phản biện và chỉnh sửa 20/3/2019 → Duyệt đăng 25/3/2019.

1. Đặt vấn đề

Thống kê là nội dung trong môn học Toán - Thống kê Y Dược (Xác suất thống kê) ở tất cả các trường đại học (ĐH) ngành Dược. Tuy nhiên, việc dạy học thống kê không chỉ dừng lại ở việc dạy cho sinh viên (SV) các khái niệm, các công thức thống kê mà còn phải dạy cho SV biết tư duy, biết vận dụng liên hệ thống kê với thực tiễn ngành nghề mà họ theo học. Để SV có tư duy, biết mở rộng và phát triển các phương pháp thống kê trong thực tiễn ngành Dược thì việc phát triển tư duy thống kê (TDTK) cho SV trong dạy học Toán - Thống kê Y Dược là một trong những nhiệm vụ quan trọng và mang lại ý nghĩa to lớn.

2. Nội dung nghiên cứu

2.1. Tư duy thống kê

TDTK chỉ xuất hiện trước một tình huống có vấn đề, hay nói cách khác TDTK bắt đầu khi có vấn đề của chủ thể cho đến khi vấn đề của chủ thể được giải quyết. “TDTK là quá trình nhận thức, phản ánh và vận dụng những thuộc tính bản chất, những mối quan hệ có tính quy luật thống kê thông qua các dữ liệu thống kê đại diện cho tổng thể cần nghiên cứu” [1].

Nhận thức là sự hiểu biết, tìm hiểu, đánh giá, tìm kiếm mối liên hệ của tập dữ liệu thống kê, nó bắt đầu từ việc xem xét tập dữ liệu một cách trực tiếp, tích cực và sáng tạo để đưa ra những nhận định, những so sánh, suy luận, những phán đoán hay kết luận có tính quy luật thống kê của tập dữ liệu.

Phản ánh là quá trình khái quát, tổng hợp, chỉ ra các thuộc tính bản chất, các đặc trưng hay tính quy luật của tập dữ liệu thống kê.

Vận dụng là quá trình vận dụng những khái niệm, những công thức, quy tắc thống kê để tiến hành phân tích, đánh giá và đưa ra các kết luận có ý nghĩa thống kê nhằm giải quyết các tình huống, các vấn đề về tập dữ liệu thống kê.

Nói đến thống kê là nói đến một ngành khoa học giúp chúng ta thực hiện các công việc như: Thu thập thông tin chân thực về đối tượng nghiên cứu với một độ tin cậy nhất định; Tổ chức và thu gọn, biểu diễn số liệu nhằm hỗ trợ quá trình tìm hiểu về vấn đề nghiên cứu hoặc đối tượng nào đó; Phân tích và rút ra những kết luận có ý nghĩa thống kê; Ứng dụng thống kê trong thực tiễn ngành nghề.

Ngày nay, thống kê là một phần quan trọng không thể thiếu đối với các nhà nghiên cứu Dược. Thống kê trở thành công cụ giúp các nhà nghiên cứu Dược học có thể ước lượng, dự đoán được hàm lượng thuốc, so sánh tác dụng của hai loại thuốc, ... Vì vậy, để đáp ứng được yêu cầu của ngành Dược thì TDTK là điều không thể thiếu đối với các nhà nghiên cứu Dược.

2.2. Các biện pháp phát triển tư duy thống kê cho sinh viên đại học ngành Dược thông qua luyện tập cho sinh viên đọc hiểu, tổ chức và thu gọn, biểu diễn dữ liệu thống kê

2.2.1. Luyện tập cho sinh viên đọc hiểu dữ liệu thống kê

a. Mục đích và ý nghĩa

Dữ liệu thống kê là một trong các bước quan trọng và cần thiết đối với việc phân tích dữ liệu thống kê. Hoạt động liên tưởng, kết nối giữa dữ liệu gốc với bảng biểu, đồ thị sẽ giúp TDTK hình thành và phát triển. Việc luyện tập cho SV đọc được dữ liệu thống kê giúp SV có khả năng hiểu được dữ liệu gốc, hiểu ý nghĩa của mẫu số liệu, từ đó phát hiện những quy luật thống kê. Với biện pháp này, TDTK của SV được phát triển thể hiện ở việc SV có khả năng đọc và hiểu các dữ liệu trên bảng biểu, đồ thị.

b. Nội dung và cách thức thực hiện

Trong chương trình môn Toán - Thống kê Y Dược, đọc hiểu dữ liệu thống kê là đọc hiểu bảng dữ liệu thống kê và biểu đồ thống kê.

- **Đối với bảng dữ liệu thống kê:** Với bảng dữ liệu, SV phải đọc được các thông số như: Kích thước mẫu, dấu hiệu nghiên cứu, đơn vị nghiên cứu và phải hiểu được vấn đề

Bảng 1: Kết quả xét nghiệm máu

Họ và tên	Địa chỉ	Tuổi	Giới tính	HC	HB	TC	BC
Phạm Mai Anh	Nam Định	30	Nữ	6,2	160	420	5,5
Đỗ Hùng	Hà Nội	40	Nam	4,5	135	300	9,2
Nguyễn Tâm Lan	Thái Bình	35	Nữ	5,0	150	350	8,0
Nguyễn Văn Lâm	Hà Nội	38	Nam	4,3	130	270	8,5
Hà Đình Thông	Hà Tây	43	Nam	5,2	155	370	6,4
Trần Thị Thu	Vĩnh Phúc	37	Nữ	4,7	150	300	6,5
Bùi Anh Tuấn	Hà Tây	45	Nam	4,8	138	320	7,0

Bảng 2: Doanh thu/tháng của 139 nhà thuốc bán lẻ ở thành phố A

Doanh thu	[15-20)	[20-25)	[25-30)	[30-35)	[35-40)	[40-45)	[45-50)	[50-55)
Số nhà thuốc	5	12	25	35	23	20	17	2

cần nghiên cứu là gì, có những thông số nào. Trước một bảng dữ liệu, SV phải tư duy, nghiên cứu và liên tưởng đến những con số thống kê.

Ví dụ 1: Cho bảng dữ liệu về kết quả xét nghiệm máu về HC (Hồng cầu), HB (Hemoglobin – Lượng huyết sắc tố trong một thể tích máu), TC (Tiểu cầu) và BC (Bạch cầu) của một số bệnh nhân như sau (xem Bảng 1):

Câu hỏi 1: Kích thước mẫu là bao nhiêu, trong đó có bao nhiêu nữ và bao nhiêu nam, các bệnh nhân có độ tuổi nằm trong khoảng nào?

Nhìn vào Bảng 1, muốn biết được kích thước mẫu, SV có thể đếm số bệnh nhân (cột Họ và tên) và biết được kích thước mẫu là 7, cột giới tính biết được 3 nữ và 4 nam. Nhìn vào cột Tuổi, SV biết được tuổi thấp nhất là 30 và cao nhất là 45 nên các bệnh nhân có độ tuổi từ 30 đến 45 tuổi.

Câu hỏi 2: Bệnh nhân nữ nào có chỉ số HC (Hồng cầu) thấp nhất? Bệnh nhân nào có chỉ số BC (Bạch cầu) cao nhất?

Bằng cách so sánh các số trong cột HC và cột Giới tính, SV biết được bệnh nhân nữ Trần Thị Thu 37 tuổi, quê ở Vĩnh Phúc có chỉ số thấp nhất là 4,7 và bệnh nhân Đỗ Hùng ở Hà Nội có chỉ số BC cao nhất là 9,2.

Qua câu hỏi 1 và 2, TDTK của SV được phát triển thông qua quá trình nhận thức và phản ánh nhằm rèn luyện cho SV nhận biết đối tượng nghiên cứu, để từ đó có thể tổng quát và đưa ra những nhận định ban đầu về tập dữ liệu.

Câu hỏi 3: Lượng Hemoglobin trung bình trong nhóm bệnh nhân trên là bao nhiêu?

Giá trị trung bình của lượng HB không xuất hiện trên bảng dữ liệu. Do vậy, để tính được giá trị trung bình của HB thì SV phải suy nghĩ, phải liên tưởng đến công thức tính giá trị trung bình của mẫu đơn giản và áp dụng công thức tính sau:

$$\bar{x}_{HB} = \frac{1}{n} \sum_{i=1}^n x_{HBi} = \frac{160+135+150+130+155+150+138}{7} = 145,43$$

Với câu hỏi 3, SV phải nhận biết về đối tượng nghiên

cứu, đồng thời tư duy để vận dụng công thức tính giá trị trung bình của loại mẫu nào để có thể tính được lượng Hemoglobin trung bình của nhóm bệnh nhân.

Câu hỏi 4: Tính giá trị trung vị của TC?

Để tính giá trị trung vị, SV phải liên tưởng đến khái niệm và cách tìm trung vị, từ đó sẽ sắp xếp các số liệu trong cột TC theo thứ tự tăng dần: 270; 300; 320; 350; 370; 420.

Muốn tìm giá trị trung vị, SV phải hiểu được giá trị trung vị là ở vị trí chính giữa của tập dữ liệu, từ đó tìm được giá trị chính giữa là 320. Khi đó đi đến kết luận giá trị trung vị là 320.

Ví dụ 2: Để ước tính doanh thu hàng tháng trung bình của các nhà thuốc bán lẻ trên địa bàn thành phố A, người ta đã tiến hành điều tra doanh thu hàng tháng của một số nhà thuốc bán lẻ trên địa bàn thành phố A, thu được kết quả (tính theo triệu đồng/tháng) như sau (xem Bảng 2):

Để SV hiểu rõ về bảng dữ liệu, giảng viên đưa ra các câu hỏi và nhận xét để SV tư duy và suy luận để trả lời câu hỏi. Chẳng hạn, giảng viên có thể đưa ra các câu hỏi đối với bảng dữ liệu trên như sau:

Câu hỏi 1: Có bao nhiêu nhà thuốc có doanh thu nhỏ hơn 20 triệu đồng/tháng và chiếm bao nhiêu phần trăm trong số các nhà thuốc được điều tra?

Với bảng dữ liệu, có thể thấy tần số của các nhà thuốc có doanh số nhỏ hơn 20 triệu đồng/tháng là 5 nhà thuốc và đếm tổng các tần số (kích thước mẫu): 139.

Khi đó, các nhà thuốc có doanh thu nhỏ hơn 20 triệu đồng/ tháng chiếm khoảng 3,6% ($\frac{5}{139} \times 100\%$).

Câu hỏi 2: Nếu coi những nhà thuốc có doanh thu nhỏ hơn 30 triệu đồng/tháng là những nhà thuốc có doanh thu thấp, tính tỉ lệ những nhà thuốc loại này?

Để tìm tỉ lệ những nhà thuốc có doanh thu thấp, cần phải tìm số các nhà thuốc có doanh thu thấp, SV sẽ đến tần số các nhà thuốc có doanh thu nhỏ hơn 30 triệu đồng/tháng được 42 nhà thuốc.

Như vậy, tỉ lệ các nhà thuốc loại này chiếm khoảng 30%.

Với câu hỏi 1 và 2, SV phải đưa ra những nhận định ban đầu về tập dữ liệu, phải xem xét, đánh giá, tổng hợp và phân tích tập dữ liệu, vận dụng các kiến thức về tỉ lệ để có thể đưa ra được các kết luận thống kê.

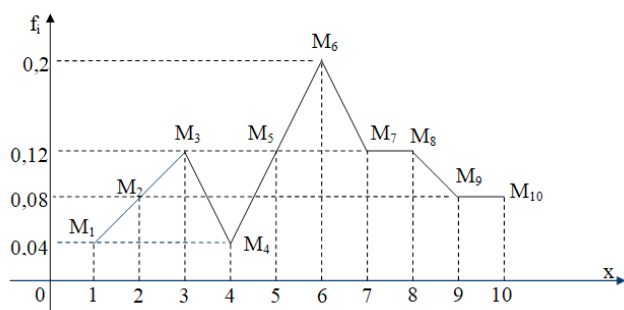
Câu hỏi 3: *Tính doanh thu trung bình/tháng của các nhà thuốc trên?*

Với việc hiểu được các số liệu trên Bảng 2, SV tìm và lựa chọn công thức tính các số đặc trưng với loại mẫu phân lớp. SV phải lấy các giá trị đại diện trong từng khoảng, để đơn giản trong việc tính toán, SV sẽ đổi biến rồi tính. Áp dụng công thức tính trung bình, chúng ta được: $\bar{x} = 3,46$.

Việc luyện tập cho SV đọc và hiểu bảng dữ liệu nhằm rèn luyện cho SV trí tưởng tượng, phép so sánh và sự tổng hợp để có thể hiểu hết được ý nghĩa của việc nghiên cứu.

- **Đối với biểu đồ thống kê:** Luyện tập cho SV nhận biết các loại biểu đồ như biểu đồ tần suất, biểu đồ tần số hay biểu đồ hình tròn. Để đọc các thông số, kích thước mẫu, các dấu hiệu quan sát được có trên biểu đồ thì SV phải suy nghĩ, tìm tòi, liên tưởng đến từng dạng biểu đồ hay nói cách khác chính là TDTK.

Ví dụ 3: Biểu đồ 1 thể hiện biểu đồ của một mẫu có kích thước là 25 (xem Biểu đồ 1):



Biểu đồ 1: Biểu đồ tần suất của mẫu (Nguồn: Trích từ [2])

Câu hỏi 1: *Dựa vào biểu đồ cho biết các giá trị có trong mẫu?*

Nhìn vào Biểu đồ 1, các giá trị của mẫu nằm ở trục hoành, tần suất tương ứng với các giá trị nằm trên trục tung, các giá trị của mẫu là 1, 2, 3, 4, 5, 6, 7, 8, 9, 10.

Câu hỏi 2: *Dựa vào biểu đồ trên, hãy cho biết giá trị mode của mẫu?*

SV phải so sánh các giá trị tần suất, giá trị có tần suất lớn nhất ứng với tần số lớn nhất để suy ra giá trị mode của mẫu nghiên cứu là 6.

Câu hỏi 3: *Hãy liệt kê dữ liệu dưới dạng bảng tần số.*

Câu hỏi này đòi hỏi mức tư duy cao hơn câu hỏi 1 và 2,

SV phải liệt kê từng cặp điểm $M_i(x_i, f_i)$, $i = 1, \dots, 10$ và với mỗi tần suất f_i tính ra tần số n_i theo công thức: $n_i = n \times f_i$.

Bảng tần số được tính ở Bảng 3 dưới đây (xem Bảng 3):

Qua rèn luyện cho SV đọc hiểu dữ liệu thống kê, gợi nhớ cho SV các bước tìm giá trị đặc trưng của mẫu đã được học ở phổ thông và trong chương Lí thuyết mẫu. Luyện tập cho SV đọc hiểu, nhận biết nhằm rèn luyện trí tưởng tượng, so sánh và sự tổng hợp để có thể đọc hiểu được số liệu trên bảng biểu, biểu đồ. Việc luyện tập cho SV đọc hiểu dữ liệu thống kê giúp cho SV đạt mức độ 1 (Biết đọc hiểu dữ liệu [1]) của TDTK mà SV ĐH ngành Dược cần đạt được. Đồng thời, SV hiểu được nguồn gốc của tập dữ liệu, từ đó có thể phân tích và xử lí được dữ liệu thống kê và đưa ra được những kết luận có ý nghĩa thống kê.

2.2.2. Luyện tập cho sinh viên tổ chức và thu gọn, biểu diễn dữ liệu thống kê

a. Mục đích và ý nghĩa

Tổ chức và thu gọn dữ liệu là một phần tất yếu trong quá trình xử lí thống kê. Trước tập dữ liệu gốc dài và có nhiều dữ liệu trùng nhau thì việc tổ chức và thu gọn là một việc làm cần thiết. Thông qua tổ chức và thu gọn mà các dữ liệu thô sẽ được tóm tắt lại bằng sử dụng các phương pháp tính toán đơn giản hoặc được trình bày trong các bảng, biểu đồ. Việc tổ chức và thu gọn dữ liệu giúp chúng ta nhanh chóng trực quan hóa và giúp chúng ta có được các mô tả ban đầu về tập dữ liệu mà ta đang quan tâm. Đồng thời, giúp chúng ta lựa chọn được phương pháp phân tích phù hợp với yêu cầu của việc nghiên cứu, từ đó có thể tìm được những thông tin mà ta cần nghiên cứu và đưa ra quyết định nhanh chóng.

b. Nội dung và cách thức thực hiện

Để giúp SV biết tổ chức và thu gọn dữ liệu thống kê thì trong quá trình giảng dạy về *Chương 1: Lí thuyết mẫu*, SV phải hiểu rõ được các cách tổ chức và thu gọn dữ liệu, trước một tập dữ liệu gốc phải biết tổ chức như thế nào và biết thu gọn theo dạng nào là phù hợp với yêu cầu nghiên cứu.

- **Thu gọn dưới dạng bảng:** Việc luyện tập cho SV thu gọn dữ liệu dưới dạng bảng giúp SV nhận biết, hiểu được tập dữ liệu, bước đầu có thể đánh giá về tập dữ liệu cần nghiên cứu.

Ví dụ 4: Để đánh giá trọng lượng của trẻ em 7 tuổi của một lớp tại một trường tiểu học có đạt trọng lượng chuẩn không, người ta tiến hành kiểm tra trọng lượng của 50 trẻ em 7 tuổi của lớp đó và kết quả thu được (tính theo kg/em) như sau (xem Bảng 4):

Với tập dữ liệu này, bước đầu SV phải đưa ra các nhận định như: mẫu có nhiều giá trị trùng nhau, các giá trị trong mẫu đều là số nguyên, kích thước mẫu là 50 nên là mẫu lớn.

Bảng 3: Bảng tần số ứng với các giá trị của mẫu

x_i	1	2	3	4	5	6	7	8	9	10
f_i	0,04	0,08	0,12	0,04	0,12	0,2	0,12	0,12	0,08	0,08
n_i	1	2	3	1	3	5	3	3	2	2

Bảng 4: Trọng lượng của 50 trẻ em 7 tuổi

20	21	24	21	24	25	20	27	30	22
23	24	25	26	27	22	24	25	23	21
20	23	25	26	23	28	29	30	23	24
25	26	27	22	24	25	21	24	25	20
28	29	30	26	22	25	26	27	28	29

Để đơn giản và tóm tắt lại mẫu dữ liệu thì SV phải tổ chức và thu gọn mẫu dữ liệu sao cho dễ nhớ, dễ hiểu và lựa chọn loại mẫu thu gọn để mô tả cho tập dữ liệu gốc. Cuối cùng, SV phải vận dụng các bước thu gọn mẫu như: liệt kê các giá trị của mẫu theo thứ tự từ bé đến lớn, đếm tần số các giá trị đó và liệt kê các giá trị ứng với tần số dưới dạng bảng, được kết quả (xem Bảng 5):

Bảng 5: Bảng mô tả mẫu thu gọn

x_1	x_1	x_2	x_3	...	x_k
n_1	n_1	n_2	n_3	...	n_k

Trong đó: $\sum_{i=1}^k n_i = n$

Tập dữ liệu trên được sắp xếp và thu gọn như sau (xem Bảng 6):

Với tập dữ liệu đã được sắp xếp giá trị từ thấp đến cao và đã được thu gọn, dễ dàng đọc được các thông số, các giá trị cần nghiên cứu, chẳng hạn có thể thấy được trọng lượng

của trẻ em 7 tuổi của lớp học đó thấp nhất là 20 kg và cao nhất là 30 kg,....

Ví dụ 5: Đo đường kính của 100 viên thuốc Tetracyclin loại 50mg/viên của một xí nghiệp sản xuất thuốc, người ta thu được kết quả như sau (xem Bảng 7):

Hãy thu gọn mẫu trên? Để đưa ra sự lựa chọn dạng mẫu thu gọn phù hợp với mẫu gốc trên, SV buộc phải tư duy và đưa ra các nhận xét về mẫu trên như: Các số liệu nhiều là số lẻ, gần nhau. Do đó, SV đưa ra kết luận lựa chọn mẫu thu gọn dạng khoảng là phù hợp.

Để thu gọn dưới dạng khoảng, cần phải thực hiện các bước sau: 1/ Chọn một khoảng [a, b) chứa tất cả các giá trị của tập dữ liệu; 2/ Chia khoảng [a, b) thành k ($2 < k < n$) khoảng con $[a, a_1)$, $[a_1, a_2)$, ..., $[a_{k-1}, b)$; 3/ Đếm tần số các khoảng con; Biểu diễn các khoảng và tần số của chúng theo bảng sau (xem Bảng 8):

Với tập dữ liệu ở Bảng 9: giá trị nhỏ nhất là 48 và giá trị lớn nhất là 51,2 nên chọn khoảng [48; 51,5) chứa 100 giá trị của tập dữ liệu và chia khoảng này thành 7 khoảng con: [48; 48,5), [48,5; 49), [49; 49,5), [49,5; 50), [50; 50,5), [50,5; 51), [51; 51,5). Sau khi thực hiện đếm tần số của các giá trị

Bảng 6: Bảng thu gọn (dạng tần số) trọng lượng của 50 trẻ em 7 tuổi

Trọng lượng (kg):	20	21	22	23	24	25	26	27	28	29	30
Số trẻ:	4	4	4	5	7	8	5	4	3	3	3

Bảng 7: Đường kính của 100 viên thuốc Tetracyclin

48,5	48,1	48,6	49,5	49,3	49,2	50	50,5	50,1	48,2
48,5	48,3	48,2	49,3	49,5	49,7	49,6	50,3	50	51,2
49,7	49,7	49,9	49,5	50,5	50,7	50,8	48,6	49,2	50,7
48,9	49	50	48,5	48,2	49,6	49,9	49,7	49,9	50,6
49,5	49,6	49,1	49,3	50,2	49,7	49,8	49,5	49,9	48,6
50,5	50,7	50,8	48,6	50	48,5	48,2	49,6	49,7	49,8
49,3	49,5	49,7	49,6	50,3	50	51,2	49,3	49,5	49,7
49,7	49,9	49,5	50,5	50,7	50,8	48,6	49,2	49,7	49,9
49,5	49,6	49,1	49,3	50,2	49,3	50,2	49,7	49,8	49,5
49,9	49,5	50,5	50,7	50,8	48,6	48,2	49,3	49,5	49,7

Bảng 8: Bảng mô tả mẫu thu gọn dạng khoảng

x_i	$[a, a_1)$	$[a_1, a_2)$...	$[a_i, a_{i+1})$...	$[a_{k-1}, b)$
n_i	n_1	n_2	...	n_{i+1}	...	n_k

Khi đó:
$$\begin{cases} \sum_{i=1}^k n_i = n \\ x_1 < x_2 < \dots < x_k \end{cases}$$

Bảng 9: Bảng thu gọn dạng khoảng đường kính của 100 viên thuốc Tetracyclin

Hàm lượng	[48-48,5)	[48,5-49)	[49-49,5)	[49,5-50)	[50-50,5)	[50,5-51)	[51-51,5)
Số viên	7	10	14	41	11	15	2

trong các khoảng trên và sắp xếp thành dạng bảng thì được mẫu thu gọn dạng khoảng như sau (xem Bảng 9):

Việc luyện tập cho SV biết tổ chức và thu gọn mẫu không chỉ giúp SV thành thạo các thao tác thống kê trong quá trình học mà đòi hỏi SV phải tư duy và đưa ra những nhận định ban đầu của mình, đưa ra cách lựa chọn, tổ chức đối với một tập dữ liệu và quyết định lựa chọn phương pháp thu gọn nào là phù hợp với yêu cầu nghiên cứu.

- Rèn luyện cho SV vẽ được biểu đồ và đồ thị: Trước khi tiến hành phân tích dữ liệu thống kê, chúng ta cần phải hiểu được tập dữ liệu, cung cấp được những tóm tắt cơ bản của tập dữ liệu thống kê đã điều tra được. Việc biểu diễn dữ liệu thống kê sẽ cung cấp những tóm tắt cơ bản. Biểu đồ là các hình vẽ hoặc đường nét hình học dùng để mô tả các tính chất, phân bố của một biến, hai biến hay nhiều biến. Biểu đồ thể hiện xu hướng, kết quả của tập dữ liệu, ... Biểu đồ đóng vai trò quan trọng trong phân tích thống kê. Để biểu diễn được biểu đồ, đòi hỏi phải tư duy sáng tạo, phải có sự tổng hợp số liệu mới cho biểu đồ chính xác. Rèn luyện cho SV cách biểu diễn và phân tích được các loại biểu đồ giúp SV có khả năng hiểu biết được bản chất, quy luật phân bố của dữ liệu, thậm chí có thể so sánh được các biến.

Có nhiều loại biểu đồ được dùng để biểu diễn các số liệu thống kê trong nghiên cứu Y Dược, mỗi loại biểu đồ biểu thị một đặc trưng riêng của số liệu thống kê, tùy thuộc vào tập số liệu và nhu cầu nghiên cứu mà lựa chọn biểu đồ thích hợp. Mỗi loại biểu đồ biểu thị những chủ đề, những tính chất, đặc trưng khác nhau, vì vậy khi vẽ biểu đồ SV phải đọc kỹ yêu cầu của đề bài để tìm hiểu các dữ liệu, thông số thể hiện trên biểu đồ, sau đó căn cứ vào yêu cầu của đề bài lựa chọn biểu đồ thích hợp để biểu diễn. Khi vẽ biểu đồ, phải đảm bảo các nguyên tắc sau: Tính khoa học (tính chính xác); Tính trực quan; Tính thẩm mỹ.

Khi vẽ biểu đồ cần lưu ý các kí hiệu sao cho dễ đọc và dễ hiểu. Giả sử ta có một mẫu thu gọn (x_1, x_2, \dots, x_k) . ($k=1, \dots, n$) với các tần số tương ứng n_1, n_2, \dots, n_k .

Bảng 11: Bảng tần suất của 38 trẻ em 3 tuổi

Trọng lượng	11	12	13	14	15	16	17	18
Số trẻ	1	2	5	7	9	6	5	3
Tần suất	0,026	0,052	0,132	0,184	0,237	0,158	0,132	0,079

Khi giảng viên rèn luyện cho SV vẽ đa giác tần suất thì giảng viên hướng dẫn SV từng bước vẽ đa giác tần suất:

Bước 1: Lập bảng tính tần suất f_i tương ứng với giá trị x_i , với $i = 1, \dots, k$.

Bước 2: Trong mặt phẳng tọa độ Oxy, xác định các điểm $M_i(x_i, f_i)$, trong đó: x_i là giá trị của mẫu, f_i là tần suất tương ứng với giá trị x_i , với $i = 1, \dots, k$.

Bước 3: Nối liên tiếp các điểm M_j .

Khi nối các điểm M_i trên mặt phẳng Oxy, chúng ta được một đường gấp khúc, đường gấp khúc gọi là đa giác tần suất của mẫu.

Ví dụ 6: Trọng lượng của 38 trẻ em 3 tuổi được cho bởi (tính theo kg) (xem Bảng 10):

Bảng 10: Bảng trọng lượng của 38 trẻ em 3 tuổi

Trọng lượng	11	12	13	14	15	16	17	18
Số trẻ	1	2	5	7	9	6	5	3

Câu hỏi: Vẽ đa giác tần suất của mẫu trên? Để giải quyết được câu hỏi này, SV phải đưa ra những nhận định ban đầu về tập dữ liệu: Là mẫu thu gọn, các giá trị của mẫu là số nguyên. Từ đó, SV lần lượt vận dụng các bước vẽ đa giác tần suất như sau:

Bước 1: Tính tần suất: $f_i = \frac{n_i}{n}$ ($i = 1, \dots, 8$).

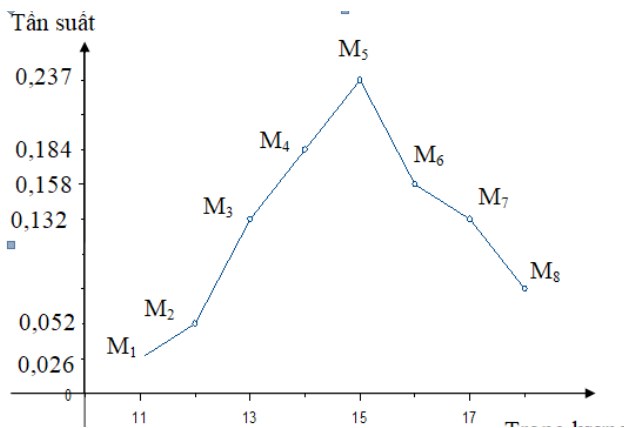
Bảng tần suất được tính ở Bảng 11:

Bước 2: Xác định các điểm $M_1(11; 0,026)$, $M_2(12; 0,052)$, $M_3(13; 0,132)$, $M_4(14; 0,184)$, $M_5(15; 0,237)$, $M_6(16; 0,158)$, $M_7(17; 0,132)$, $M_8(18; 0,079)$ trên trục tọa độ Oxy.

Vẽ trục tọa độ Oxy và lấy các điểm $M_i(x_i, f_i)$ trên mặt phẳng tọa độ.

Bước 3: Nối liên tiếp các điểm M_i ($i = 1, \dots, 8$) với nhau, ta được đa giác tần suất.

Giảng viên có thể gọi SV lên bảng vẽ biểu đồ và chữa (xem Biểu đồ 2).



Biểu đồ 2: Đa giác tần suất về trọng lượng của 38 trẻ em

Nếu X là biến ngẫu nhiên liên tục với hàm mật độ $p(x)$ chưa biết, để rèn luyện cho SV vẽ được biểu đồ thì giảng hướng dẫn sinh lần lượt thực hiện các bước sau:

Bước 1: Chia mẫu thành k khoảng con $[\alpha_{i+1} - \alpha_i)$ bằng nhau với độ dài mỗi khoảng là $h = \alpha_{i+1} - \alpha_i$ và đếm tần số của mỗi khoảng là n_i ($i = 1, 2, \dots, k$).

Lập bảng tính tần số f_i và $\frac{f_i}{h}$.

Bước 2: Trong mặt phẳng tọa độ Oxy xác định các giá trị trên các trục tọa độ:

Trục hoành chứa các giá trị là đầu mút của các khoảng $[\alpha_{i+1} - \alpha_i)$.

Trục tung chứa các giá trị $\frac{f_i}{h}$.

Bước 3: Vẽ các hình chữ nhật có đáy là $[\alpha_i, \alpha_{i+1}]$ và chiều cao là $\frac{f_i}{h}$.

3. Kết luận

Phát triển TDTK cho SV ĐH ngành Dược thông qua việc luyện tập cho SV đọc hiểu và biết tổ chức, thu dọn dữ liệu, biểu diễn dữ liệu không những giúp SV hiểu được ý nghĩa của thống kê, học tập tốt hơn môn học Toán - Thống kê Y Dược mà còn giúp SV ngành Dược đưa ra được những quyết định hiệu quả hơn về nghề nghiệp khi họ học thống kê.

Tài liệu tham khảo

- [1] Quách Thị Sen, (2019), *Phát triển tư duy thống kê cho sinh viên trong dạy học Toán - Thống kê Y Dược ở các trường đại học ngành Dược*, Tạp chí Khoa học Giáo dục, số 14, tr. 29.
- [2] Nguyễn Phan Dũng, Quách Thị Sen, Phạm Thị Hồng Cẩm, (2018), *Xác suất và thống kê*, NXB Y học.
- [3] Trần Đức Chiên, (2007), *Rèn luyện tư duy thống kê cho học sinh trong dạy học thống kê - xác suất ở môn Toán trung học phổ thông*, Luận án Tiến sĩ Giáo dục học, Viện Chiến lược và Chương trình Giáo dục.
- [4] Hoàng Nam Hải, (2013), *Phát triển năng lực suy luận thống kê cho sinh viên cao đẳng chuyên nghiệp*, Luận án Tiến sĩ Giáo dục học, Trường Đại học Vinh.
- [5] Trần Ngọc Lan, Trương Thị Tố Mai, (2007), *Rèn luyện tư duy cho học sinh trong dạy học toán bậc Tiểu học*, NXB Trẻ.
- [6] Chu Cẩm Thơ, (2015), *Phát triển tư duy thông qua dạy học môn Toán ở trường phổ thông*, NXB Đại học Sư phạm Hà Nội.
- [7] Randall E. Groth, (2003), *Development of a high school statistical thinking framework*, Luận án tiến sĩ, Illinois State U., USA.
- [8] Snee R., (1999), *Discussion: Development and use of statistical thinking: A new era*, International Statistical Review, 67(3), 255-258.

DEVELOPING STATISTICAL THINKING FOR UNIVERSITY STUDENTS IN PHARMACY THROUGH PRACTICE FOR STUDENTS TO READ, ORGANIZE, REDUCE, AND REPRESENT STATISTICAL DATA

Quach Thi Sen

Hanoi University of Pharmacy
13 - 15 Le Thanh Tong, Hoan Kiem, Hanoi, Vietnam
Email: senqtdhd@gmail.com

ABSTRACT: *Developing statistical thinking is one of the important tasks for training university students in Pharmacy schools. Through the concept of statistical thinking and the levels of statistical thinking that university students in Pharmacy need to achieve, the paper proposes measures to develop statistical thinking in teaching Math - medicine statistics for university students in Pharmacy, practice for students to read statistical data as well as organize, reduce, and represent the statistical data.*

KEYWORDS: Statistical thinking; data; practice; Pharmacy; students.