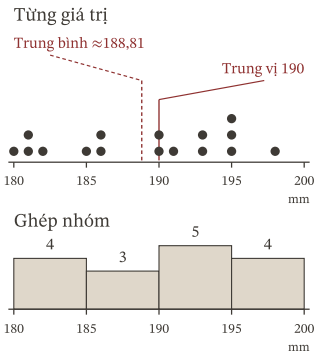
ZO MATH

ÔN THI TOÁN THPT 2027

Đọc dữ liệu và đo xu thế trung tâm

Lời giải

Lời giải, đáp án và hướng dẫn đối chiếu.



Mở trang Đọc dữ liệu và đo xu thế trung tâm

R2-G01

Phiên bản nội dung 0.4 · Ngày tạo PDF: 04/11/2026

© 2026 ZO Math. Bản quyền nội dung thuộc ZO Math.

Mục lục

Lời giải **2**
Bài học — Khởi động . . . 2
Bài học — Các nhiệm vụ trong mạch học . . . 2
Luyện tập — Chặng 1 . . . 3
Luyện tập — Chặng 2 . . . 5
Luyện tập — Chặng 3 . . . 6
Kiểm tra — Chỉ đối chiếu sau khi kết thúc lượt . . . 8
Ôn lại . . . 10
Ôn nền . . . 16
Sửa lỗi — Đối chiếu đúng ý đã tự làm . . . 17
Nguồn và cách dùng dữ liệu **19**
Tài liệu toán học . . . 20
Palmer Penguins . . . 21
Bike Sharing . . . 21
NOAA/NCEI . . . 21

Lời giải

Giữ bài làm ban đầu trước khi mở phần tương ứng. Các ví dụ trong Bài học đã được giải tại chỗ; ở đây là lời giải của những nhiệm vụ tự làm. Lời giải Kiểm tra chỉ dùng sau khi kết thúc lượt.

Bài học — Khởi động

Lời giải Nhiệm vụ 01 — Ba thao tác nền

a) Ba số đã theo thứ tự không giảm. Trung bình là $\frac{181+186+195}{3} = \frac{562}{3} \approx 187{,}3$ mm.

b) $\frac{3}{12} \cdot 100\% = 25\%$. Trong nhóm giả định này, 25% cá thể có chiều dài cánh bơi từ 190 mm trở lên.

c) $190 \in [190; 195)$ vì $190 \le 190 < 195$. Khoảng $[185; 190)$ không nhận đầu mút phải 190.

Xem đề bài (mở trên trang học liệu).

Bài học — Các nhiệm vụ trong mạch học

Lời giải Nhiệm vụ 02 — Đếm đúng mẫu dùng để tính

a) Đơn vị quan sát là một cá thể chim; biến là chiều dài cánh bơi; đơn vị đo mm.

b) Có 16 giá trị hợp lệ nên phải chia cho 16: $\overline{x} = \frac{3021}{16} = 188{,}8125$ mm. Số 17 là số dòng trong bảng, gồm cả dòng thiếu số đo chiều dài.

c) Sửa thành: "Dòng 4 chưa có số đo chiều dài cánh bơi." Không có căn cứ gán chiều dài bằng 0.

Xem đề bài (mở trên trang học liệu).

Lời giải Nhiệm vụ 03 — Vị trí và giá trị

a) 4,5 diễn tả vị trí nằm giữa vị trí 4 và 5, không phải chiều dài. Hai giá trị ở đó là 182 và 185 mm, nên $Q_1 = \frac{182+185}{2} = 183{,}5$ mm.

b) Không. Trong mẫu năm số đo ở Ví dụ 02, quan sát 190 mm ở vị trí thứ ba dùng để xác định trung vị Q_2. Theo cách xác định tứ phân vị bằng trung vị của hai nửa đã nêu trong Bài học, nửa dưới chỉ gồm hai quan sát trước vị trí ấy, nửa trên chỉ gồm hai quan sát sau vị trí ấy; mẫu ban đầu vẫn giữ đủ năm số đo. Đưa nó vào cả hai nửa là dùng quy ước khác, có thể làm thay đổi kết quả. Cần nêu và dùng nhất quán quy ước, không đổi theo đáp số mong muốn.

c) Mẫu ban đầu có tổng 15, trung bình 3 phút và trung vị 2 phút. Tổng mới là 35, trung bình 7 phút; trung vị vẫn bằng 2 phút. Giá trị cuối tăng 20 làm trung bình tăng $\frac{20}{5} = 4$ phút, nhưng giá trị ở vị trí thứ ba chưa đổi. Với câu hỏi về người ở giữa theo thời gian chờ, dùng trung vị 2 phút. Kết luận này không cho phép quên người chờ 27 phút khi đánh giá toàn bộ tình huống.

Xem đề bài (mở trên trang học liệu).

Lời giải Nhiệm vụ 04 — Thông tin còn lại sau ghép nhóm

a) Có $4 + 3 + 5 = 12$ giá trị nhỏ hơn 195 mm, tức là 12 cá thể trong mẫu có chiều dài cánh bơi dưới ngưỡng này.

b) Không. Bảng chỉ cho biết năm giá trị thuộc $[190; 195)$; trung điểm 192,5 là giá trị đại diện dùng trong phép ước lượng.

c) Nếu chỉ mắc lỗi đếm 195 hai lần, tổng tần số sẽ vượt 16. Kiểm thêm quy tắc đầu mút để biết phải sửa ở nhóm nào, thay vì chỉ ép tổng về đúng số.

Xem đề bài (mở trên trang học liệu).

Lời giải Nhiệm vụ 05 — Tự nội suy và giải thích kết quả

a) Q_3 ứng với mức 75% tổng tần số, nên $r = 3 \cdot \frac{49}{4} = 36{,}75$. Mức này nằm giữa tần số tích lũy 30 trước nhóm $[190; 195)$ và 40 đến hết nhóm. Vì vậy chọn nhóm ấy, với $a = 190$ mm, $F = 30$, $f = 10$, $h = 5$ mm:

$$\begin{aligned} Q_{3,\text{nhóm}} &= 190 + \frac{36{,}75 - 30}{10} \cdot 5 \\ &= 193{,}375 \\ &\approx 193{,}38 \text{ mm}. \end{aligned}$$

Từ 30 đến 36,75 còn thiếu 6,75, chiếm $\frac{6{,}75}{10} = 0{,}675$ tần số nhóm. Giả sử tần số tích lũy tăng đều, ta lấy 0,675 độ rộng 5 mm, được 3,375 mm; cộng vào 190 mm được 193,375 mm. Đây là giá trị ước lượng tại mức tích lũy 75%, không phải số thứ tự của một cá thể.

b) Phép tính 187,25 đúng cho trung vị của mẫu ghép nhóm và được dùng làm ước lượng. Nó không khôi phục trung vị gốc 186 mm. Sửa câu thành: “Từ bảng nhóm này, trung vị được ước lượng là 187,25 mm; tính trực tiếp trên mẫu gốc cho 186 mm.”

Xem đề bài (mở trên trang học liệu).

Lời giải Nhiệm vụ 06 — Sửa câu kết luận

a) “Tháng 01/2012, hệ thống Capital Bikeshare ghi nhận trung bình khoảng 3121 lượt thuê mỗi ngày.” Không đổi lượt thuê thành người khác nhau.

b) “Tháng 01/2023 tại trạm Seattle–Tacoma có 13 ngày được ghi lượng giáng thủy bằng 0 mm; trong đó 6 ngày có ký hiệu vết giáng thủy.” Ký hiệu T bác bỏ cách hiểu rằng mọi số 0 đều xác nhận hoàn toàn không có giáng thủy.

c) Chưa đủ căn cứ. Một tháng tại một trạm chỉ mô tả đúng phạm vi đó. Muốn kết luận về khí hậu lâu dài, cần dữ liệu của nhiều năm, kiểm tra xem cách đo qua các năm có so sánh được với nhau hay không và các trạm đo có đủ để nói về khu vực muốn xét hay không. Riêng bảng này chưa đủ.

Xem đề bài (mở trên trang học liệu).

Luyện tập — Chặng 1

Lời giải Nhiệm vụ 07 — Mẫu Adélie năm 2008

a) Một quan sát ứng với một cá thể Adélie; biến chiều dài cánh bơi, đơn vị mm; $n = 12$.

b) Dãy đã sắp:

$$\begin{aligned} &181;\ 185;\ 186;\ 186;\ 187;\ 188;\\ &190;\ 191;\ 193;\ 194;\ 195;\ 200. \end{aligned}$$

Tổng 2276 mm nên $\overline{x} = \frac{2276}{12} \approx 189{,}67$ mm. Hai vị trí giữa là 6 và 7, mang giá trị 188 và 190, nên $M_e = 189$ mm.

Trong nhóm 12 cá thể Adélie năm 2008 này, có sáu số đo nhỏ hơn 189 mm và sáu số đo lớn hơn 189 mm.

Mốt là 186 mm, xuất hiện hai lần; các giá trị còn lại xuất hiện một lần. Trung vị 189 mm nằm giữa hai giá trị chính giữa của mẫu đã sắp, không phải một chiều dài bắt buộc xuất hiện trong dữ liệu.

c) Trong mẫu 13 giá trị, 185 và 186 đều xuất hiện hai lần; các giá trị khác xuất hiện một lần. Mẫu có hai mốt 185 và 186 mm. Đây là mẫu đã thêm một quan sát, không phải đổi đáp án của mẫu 12 giá trị ở ý b.

Xem đề bài (mở trên trang học liệu).

Lời giải Nhiệm vụ 08 — Tứ phân vị khi số quan sát thay đổi

a) Gọi x_i là số đo ở vị trí thứ i trong dãy đã sắp, giữ đủ các số đo trùng nhau. Với 12 giá trị, nửa dưới gồm x_1 đến x_6, nửa trên gồm x_7 đến x_{12}. Trung vị của mỗi nửa được lấy từ hai giá trị giữa của nửa ấy:

$$Q_1 = \frac{x_3 + x_4}{2} = \frac{186 + 186}{2} = 186;$$
$$Q_2 = \frac{x_6 + x_7}{2} = \frac{188 + 190}{2} = 189;$$
$$Q_3 = \frac{x_9 + x_{10}}{2} = \frac{193 + 194}{2} = 193{,}5.$$

b) Chỉ xét 11 số đo đầu, không đưa số đo cuối 195 mm của danh sách ban đầu vào mẫu này, dãy 11 giá trị đã sắp là

$$181;\ 185;\ 186;\ 186;\ 187;\ 188;$$
$$190;\ 191;\ 193;\ 194;\ 200.$$

$Q_2 = x_6 = 188$ mm. Số đo ở vị trí thứ sáu dùng để xác định trung vị của mẫu 11 giá trị. Nửa dưới gồm năm số đứng trước vị trí ấy; nửa trên gồm năm số đứng sau. Quan sát thứ sáu vẫn thuộc mẫu ban đầu, nhưng không nằm trong hai nửa dùng để tìm Q_1 và Q_3. Kết quả là $Q_1 = 186$, $Q_3 = 193$ mm.

c) Hai ý xét hai mẫu khác nhau. Giảm số quan sát làm thay đổi các vị trí và có thể làm thay đổi các số đặc trưng; không thể dùng đáp án của mẫu 11 giá trị thay cho mẫu 12 giá trị.

d) Mỗi mẫu có tổng 72, nên trung bình bằng 9; hai giá trị ở giữa là 8 và 10, nên trung vị cũng bằng 9. Nửa dưới A là 2; 4; 6; 8, nửa trên là 10; 12; 14; 16, cho $Q_1 = 5$, $Q_3 = 13$. Nửa dưới B là 2; 8; 8; 8, nửa trên là 10; 10; 10; 16, cho $Q_1 = 8$, $Q_3 = 10$. B có sáu giá trị tập trung tại 8 và 10, còn A có các giá trị tăng đều từ 2 đến 16. Hai tứ phân vị cho thấy sự khác nhau giữa các nửa mẫu mà trung bình và trung vị chung chưa cho thấy.

Xem đề bài (mở trên trang học liệu).

Luyện tập — Chặng 2

Lời giải Nhiệm vụ 09 — Tự lập bảng

Khoảng, mm	Tần số	Trung điểm, mm	Tần số tích lũy
[180; 185)	1	182,5	1
[185; 190)	5	187,5	6
[190; 195)	4	192,5	10
[195; 200)	1	197,5	11
[200; 205)	1	202,5	12

190 thuộc [190; 195), 200 thuộc [200; 205). Tổng tần số 12 khớp cỡ mẫu.

Bảng đếm số cá thể theo từng khoảng chiều dài cánh bơi, không ghi riêng chiều dài của từng cá thể. Trung điểm chỉ là đại diện của khoảng. Chỉ biết khoảng chứa một số đo thì chưa biết số đo ấy chính xác là bao nhiêu.

d) Một cách tạo là A: 0; 0; 4; 4, B: 3; 3; 7; 7. Cả hai có tần số 2; 2, nhưng trung bình lần lượt bằng 2 và 5. Từ hai trung điểm 2 và 6, trung bình ghép nhóm bằng $\frac{2\cdot 2+2\cdot 6}{4}=4$. Bảng không còn biết mỗi quan sát nằm ở đâu trong khoảng; vì thế cùng kết quả 4 không thể khôi phục cả hai trung bình khác nhau. Có nhiều cặp mẫu khác cũng đúng; cần kiểm điều kiện mỗi số thuộc đúng khoảng và hai trung bình thực sự khác nhau.

Xem đề bài (mở trên trang học liệu).

Lời giải Nhiệm vụ 10 — Ước lượng trên mẫu lớn hơn

a) Tần số tích lũy lần lượt là 0; 1; 4; 19; 37; 46; 48; 50; 50. Tổng các tích tần số với trung điểm là 9600, nên $\overline{x}_{\text{nhóm}}=\frac{9600}{50}=192$ mm.

b) Các phép nội suy:

Số cần tìm	Mức tần số tích lũy r	Nhóm, mm	F	f	Kết quả, mm
Q_1	12,5	[185; 190)	4	15	187,83
$M_e=Q_2$	25	[190; 195)	19	18	191,67
Q_3	37,5	[195; 200)	37	9	195,28

Chẳng hạn, để tìm Q_1, mức tích lũy cần đạt là 12,5. Mức này nằm giữa 4 và 19, nên chọn [185; 190). Cần thêm $12{,}5-4=8{,}5$ trong tổng tần số 15 của nhóm. Theo giả thiết tích lũy tăng đều, lấy $\frac{8{,}5}{15}$ độ rộng 5 mm rồi cộng vào 185 mm, được $Q_{1,\text{nhóm}}\approx 187{,}83$ mm. Hai phép tính còn lại cùng cách: trung vị là $190+\frac{25-19}{18}\cdot 5$; Q_3 là $195+\frac{37{,}5-37}{9}\cdot 5$.

c) Các nhóm đều rộng 5 mm. Nhóm có tần số lớn nhất duy nhất là [190; 195), tần số 18, hai bên là 15 và 9:

$$M_{o,\text{nhóm}}=190+\frac{18-15}{(18-15)+(18-9)}\cdot 5$$
$$=191{,}25\text{ mm}.$$

d) Các kết quả trên mô tả bảng ghép nhóm của 50 số đo cánh bơi Adélie năm 2008. Bảng không giữ từng chiều dài. Dùng trung điểm và nội suy tạo các số đặc trưng dùng để ước lượng. Khác biệt với số gốc không tự chứng minh có lỗi tính toán.

Xem đề bài (mở trên trang học liệu).

Lời giải Nhiệm vụ 11 — Độ rộng nhóm có quyết định mọi sai khác không?

a) Các trung điểm là 175; 185; 195; 205:

$$\begin{aligned}\overline{x}_{\text{nhóm}} &= \frac{\begin{aligned}&1\cdot 175 + 18\cdot 185\\ &+ 27\cdot 195 + 4\cdot 205\end{aligned}}{50}\\ &= \frac{9590}{50}\\ &= 191{,}8 \text{ mm}.\end{aligned}$$

b) Bảng rộng 10 mm cho 191,8, cách trung bình gốc 191,04 một lượng 0,76 mm. Bảng rộng 5 mm cho 192, cách trung bình gốc 0,96 mm. Với đúng 50 số đo cánh bơi đang xét, kết quả từ nhóm rộng 10 mm gần trung bình gốc hơn; các số đo không thay đổi, chỉ cách chia khoảng thay đổi.

c) Phát biểu "chắc chắn" là sai; chính mẫu đang xét là một phản ví dụ đối với số trung bình. Khoảng hẹp hơn giúp xác định số đo nằm trong một phạm vi hẹp hơn. Tuy nhiên, kết quả ước lượng gần số tính từ mẫu gốc đến đâu còn phụ thuộc từng số đo nằm ở đâu trong khoảng và cách chọn đầu mút các khoảng. Không đảo thành quy tắc "nhóm rộng hơn luôn tốt hơn".

Xem đề bài (mở trên trang học liệu).

Luyện tập — Chặng 3

Lời giải Nhiệm vụ 12 — Chọn số cho bản tóm tắt lượt thuê

a) $\overline{x} = \frac{103137}{29} \approx 3556{,}45$ lượt/ngày. Số trung bình phù hợp với câu hỏi tổng lượt thuê chia đều cho mỗi ngày là bao nhiêu, vì tổng bằng số ngày nhân với trung bình chưa làm tròn.

Có thể viết: "Trong tháng 02/2012, Capital Bikeshare ghi nhận trung bình khoảng 3556,45 lượt thuê mỗi ngày."

b) Chọn trung vị 3777 lượt mỗi ngày. Khi xếp 29 số lượt thuê từ nhỏ đến lớn, vị trí chính giữa là thứ 15, và giá trị ở vị trí ấy là 3777 lượt. Số 15 là số thứ tự trong dãy đã sắp, không phải ngày 15 của tháng; 3777 mới là số lượt thuê cần báo cáo. Trung vị không được tính bằng cách chia tổng, nên không thể lấy $29 \cdot 3777$ để khôi phục tổng lượt thuê trong tháng.

c) Mẫu có 29 giá trị khác nhau; mỗi giá trị có tần số 1. Các tần số bằng nhau, nên mẫu gốc không có mốt theo quy ước đã học. Lượt thuê không phải người dùng khác nhau; một người có thể thực hiện nhiều lượt.

d) Mẫu giả định có tổng 100 lượt, trung bình 20 lượt mỗi ngày và trung vị 12 lượt. Dãy 10; 12; 12; 14; 52 đã sắp từ nhỏ đến lớn, nên ngày đứng thứ ba theo mức sử dụng có 12 lượt. Người viết bản tin chọn thông tin này để mô tả mức sử dụng ở giữa dãy. Người phụ trách vận hành cần tổng 100 lượt và mức chia đều 20 lượt mỗi ngày.

Bốn ngày có từ 10 đến 14 lượt, còn một ngày có 52 lượt. Nếu chỉ báo cáo trung vị 12 lượt, người đọc chưa biết ngày có mức sử dụng cao nhất và tổng lượt của cả năm ngày. Ngày có 52 lượt có thể quan trọng đối với việc bố trí phục vụ, dù nó không làm thay đổi giá trị ở giữa dãy; dữ liệu này cũng chưa đủ để suy nhu cầu của mọi ngày tương lai.

Xem đề bài (mở trên trang học liệu).

Lời giải Nhiệm vụ 13 — Trung bình của các nhóm

a) Trung bình nhóm ngày làm việc là $\frac{77403}{20} = 3870{,}15$ lượt/ngày. Trung bình những ngày còn lại là $\frac{25734}{9} \approx 2859{,}33$ lượt/ngày.

b) Mỗi ngày là một quan sát. Trung bình chung:

$$\begin{aligned} \overline{x} &= \frac{77403 + 25734}{20 + 9} \\ &= \frac{103137}{29} \\ &\approx 3556{,}45 \text{ lượt/ngày}. \end{aligned}$$

Nếu dùng hai trung bình nhóm chưa làm tròn, nhân chúng lần lượt với 20 và 9 sẽ khôi phục hai tổng 77403 và 25734. Cộng hai tổng rồi chia cho 29 cho kết quả trên. Còn chia đôi tổng hai trung bình cho khoảng 3364,74 lượt/ngày: cách ấy tính hai trung bình ngang nhau dù chúng lần lượt đại diện cho 20 ngày và 9 ngày. Trong cả tháng, hai nhóm chiếm $\frac{20}{29}$ và $\frac{9}{29}$ số ngày. Nhân từng trung bình với đúng số ngày của nhóm rồi cộng lại mới tính đủ đóng góp của từng ngày vào tổng.

c) Có thể kết luận: “Trong dữ liệu Capital Bikeshare tháng 02/2012, trung bình lượt thuê của nhóm ngày làm việc cao hơn nhóm còn lại.” Số lượt thuê còn có thể chịu ảnh hưởng của thời tiết, sự kiện và các yếu tố khác. Riêng bảng này chưa cho biết có phải vì là ngày làm việc mà số lượt thuê cao hơn hay không.

Xem đề bài (mở trên trang học liệu).

Lời giải Nhiệm vụ 14 — Trung vị bằng 0

a) Tổng là 2,3 mm nên trung bình $\frac{2{,}3}{31} \approx 0{,}07$ mm/ngày. Khi sắp 31 giá trị, giá trị ở vị trí thứ 16 bằng 0, nên trung vị bằng 0 mm.

b) Sửa thành: “Trung vị bằng 0 mm, nhưng tổng các giá trị lượng giáng thủy được ghi trong tháng là 2,3 mm.” Trung vị không tính bằng tổng chia cho cỡ mẫu, cũng không quyết định tổng.

c) Có thể viết: “Tháng 07/2023 tại trạm Seattle–Tacoma có 30 ngày có lượng giáng thủy được ghi bằng 0 và ngày 24/07 ghi 2,3 mm; trung vị bằng 0 mm.” Câu này giữ cả hiện tượng phổ biến và ngày có giá trị dương. Chỉ cần chọn những thông tin giúp người đọc hiểu đặc điểm của tháng đang xét; không cần đưa tất cả các số đặc trưng vào một câu.

Xem đề bài (mở trên trang học liệu).

Kiểm tra — Chỉ đối chiếu sau khi kết thúc lượt

Lời giải Nhiệm vụ 15 — Đọc, tính và ghép nhóm một mẫu mới

a) Mỗi quan sát là một cá thể chim cánh cụt Adélie trong phần dữ liệu năm 2009; biến là chiều dài cánh bơi, đơn vị mm, cỡ mẫu 12.

b) Dãy đã sắp:

$$181;\ 183;\ 184;\ 190;\ 190;\ 191;$$
$$192;\ 193;\ 197;\ 198;\ 199;\ 203.$$

Tổng 2301 mm nên $\overline{x} = \frac{2301}{12} = 191{,}75$ mm. Gọi x_i là số đo ở vị trí thứ i trong dãy đã sắp. Mỗi nửa có sáu giá trị, nên Q_1 dùng x_3, x_4, còn Q_3 dùng x_9, x_{10}:

$$Q_1 = \frac{x_3 + x_4}{2} = \frac{184 + 190}{2} = 187;$$
$$M_e = Q_2 = \frac{x_6 + x_7}{2} = \frac{191 + 192}{2} = 191{,}5;$$
$$Q_3 = \frac{x_9 + x_{10}}{2} = \frac{197 + 198}{2} = 197{,}5.$$

Các kết quả mô tả chiều dài cánh bơi của đúng nhóm mười hai cá thể đang xét; đơn vị là mm. Mốt là 190 mm: giá trị này xuất hiện hai lần, còn mỗi giá trị khác xuất hiện một lần.

c) Tần số của các khoảng từ $[180;185)$ đến $[200;205)$ lần lượt là $3; 0; 5; 3; 1$, tổng 12. Khoảng $[185;190)$ không chứa giá trị 190.

Xem đề bài (mở trên trang học liệu).

Lời giải Nhiệm vụ 16 — Bảng ghép nhóm Adélie 2009

a) Tần số tích lũy là $0; 1; 5; 16; 35; 47; 51; 51; 52$. Tổng các tích tần số với trung điểm là 10020, nên

$$\overline{x}_{\text{nhóm}} = \frac{10020}{52} \approx 192{,}69 \text{ mm}.$$

Trung bình này được tính từ bảng ghép nhóm của 52 số đo cánh bơi Adélie năm 2009.

b) Các nhóm rộng $h = 5$ mm. Trong bảng dưới, r là mức tần số tích lũy cần đạt; F là mức trước nhóm được chọn, còn f là tần số của nhóm ấy:

Số cần tìm	r	Nhóm, mm	a	F	f	Kết quả, mm
Q_1	13	$[185;190)$	185	5	11	188,64
$M_e = Q_2$	26	$[190;195)$	190	16	19	192,63
Q_3	39	$[195;200)$	195	35	12	196,67

Với trung vị, $r = \frac{52}{2} = 26$ nằm giữa mức tích lũy 16 và 35, nên dùng $[190;195)$. Phần còn thiếu là 10 trong tần số nhóm 19. Theo giả thiết tăng đều, giá trị ước lượng là $190 + \frac{10}{19} \cdot 5 \approx 192{,}63$ mm. Tương tự, hai tứ phân vị được tính bằng $185 + \frac{13-5}{11} \cdot 5$ và $195 + \frac{39-35}{12} \cdot 5$.

c) Các nhóm rộng bằng nhau. Nhóm $[190; 195)$ có tần số cao nhất duy nhất 19, hai bên là 11 và 12:

$$M_{o,\text{nhóm}} = 190 + \frac{19-11}{(19-11)+(19-12)} \cdot 5 \approx 192{,}67 \text{ mm}.$$

d) 210 thuộc $[210; 215)$, không thuộc $[205; 210)$. Trung vị ở ý b được tính từ mẫu ghép nhóm và là ước lượng cho trung vị gốc; chỉ bảng nhóm không cho phép khẳng định bằng nhau chính xác.

Xem đề bài (mở trên trang học liệu).

Lời giải Nhiệm vụ 17 — Số nào phù hợp với báo cáo?

a) $\overline{x} = \frac{164875}{31} \approx 5318{,}55$ lượt/ngày. Có thể viết: "Trong tháng 03/2012, Capital Bikeshare ghi nhận trung bình khoảng 5318,55 lượt thuê mỗi ngày." Số trung bình cho biết tổng lượt thuê chia đều cho mỗi ngày là bao nhiêu.

b) Chọn trung vị 5382 lượt mỗi ngày. Trong dãy 31 số lượt thuê đã sắp từ nhỏ đến lớn, giá trị thứ 16 bằng 5382; đó là số lượt thuê ở vị trí chính giữa. Ý a hỏi mức chia đều tổng lượt thuê cho 31 ngày, còn ý b hỏi giá trị ở giữa dãy đã sắp. Vì cách xác định khác nhau, hai số trả lời hai câu hỏi khác nhau.

c) Mẫu có 31 giá trị khác nhau, mỗi giá trị xuất hiện một lần. Theo quy ước đã học, mẫu không có mốt vì các tần số đều bằng nhau. Chưa thể suy ra số người dùng khác nhau, hoặc khẳng định một ngày tương lai sẽ có đúng số lượt thuê bằng trung bình này. Một người có thể thuê nhiều lượt, còn số lượt thuê của một ngày tương lai không được xác định bởi trung bình tháng đã qua. Em chỉ cần nêu một điều chưa thể kết luận và giải thích vì sao dữ liệu này chưa đủ.

d) Một giá trị tăng thêm 30 phút trong mẫu 6 giá trị nên trung bình tăng $\frac{30}{6} = 5$ phút. Trung bình ban đầu bằng $\frac{35}{6} \approx 5{,}83$ phút, sau thay đổi bằng $\frac{65}{6} \approx 10{,}83$ phút. Cả hai dãy đã sắp có hai giá trị ở giữa là 4 và 5, nên trung vị vẫn bằng 4,5 phút. Tổng tăng từ 35 lên 65 phút; trung vị giữ nguyên không có nghĩa tổng giữ nguyên. Khi hỏi về tổng thời gian, phải ghi nhận thay đổi này.

Xem đề bài (mở trên trang học liệu).

Lời giải Nhiệm vụ 18 — Đọc một tháng quan trắc

a) Trung bình $\frac{147}{30} = 4{,}9$ mm/ngày. Khi sắp mẫu, giá trị ở vị trí thứ 15 là 0, ở vị trí thứ 16 là 1,3. Vì vậy

$$M_e = \frac{0 + 1{,}3}{2} = 0{,}65 \text{ mm}.$$

Trung bình và trung vị trên mô tả lượng giáng thủy được ghi theo ngày tại trạm Seattle–Tacoma trong tháng 11/2023. Có 15 số 0 trong 30 giá trị không làm trung vị tự động bằng 0.

b) "Có 15 ngày có lượng giáng thủy được ghi bằng 0; trong đó ngày 07/11 có ký hiệu vết giáng thủy." Một số 0 được ghi trong bảng chưa đủ để khẳng định hoàn toàn không có giáng thủy; cần đọc cả ký hiệu đi kèm.

c) Một tháng ở một trạm chưa đủ để suy ra xu thế dài hạn. Cần dữ liệu của nhiều năm, kiểm tra xem cách đo có so sánh được giữa các thời điểm hay không và các trạm đo có đủ để nói về khu vực muốn xét hay không.

Xem đề bài (mở trên trang học liệu).

Ôn lại

Lời giải Nhiệm vụ 19 — Những con số đang mô tả điều gì?

Câu 1

a) Đối tượng được đo là các chim Chinstrap trong phần trích năm 2007; mỗi số là chiều dài cánh bơi, đơn vị mm. Có 12 giá trị. Dãy đã sắp:

$$178;\ 188;\ 192;\ 193;\ 193;\ 194;$$
$$195;\ 196;\ 197;\ 197;\ 198;\ 198.$$

b) Tổng là 2319 mm nên số trung bình bằng 193,25 mm. Hai vị trí giữa là thứ 6 và thứ 7, mang các giá trị 194 và 195 mm, cho trung vị 194,5 mm. Hai nửa có sáu số mỗi nửa; $Q_1 = 192{,}5$ mm và $Q_3 = 197$ mm.

c) Có ba mốt 193, 197 và 198 mm: mỗi giá trị xuất hiện hai lần, còn mỗi giá trị khác xuất hiện một lần. Chỉ ghi một mốt sẽ bỏ sót hai giá trị cùng đạt tần số lớn nhất.

d) Trong 12 quan sát Chinstrap được ghi nhận năm 2007 này, sáu chiều dài nhỏ hơn và sáu chiều dài lớn hơn 194,5 mm. Trung vị không nhất thiết là một chiều dài đã quan sát.

Câu 2

a) Giá trị 195 thuộc $[195; 200)$; trung điểm của $[190; 195)$ là 192,5 mm.

b) Có hai số đo nhỏ hơn 190 mm, nằm trong các khoảng đứng trước khoảng đang xét. Đây là số giá trị đã được đếm trước khi đến khoảng từ 190 mm đến dưới 195 mm.

c) Tổng các tích trung điểm với tần số là 2320:

$$\overline{x}_{\text{nhóm}} = \frac{2320}{12} \approx 193{,}33 \text{ mm}.$$

Kết quả khác 193,25 mm vì mỗi giá trị đã được thay bằng trung điểm nhóm. Sự khác nhau này không tự chứng minh bảng sai.

d) Không thể khôi phục chính xác từng số đo từ bảng. Chẳng hạn, bảng chỉ cho biết có 6 cá thể trong khoảng $[195; 200)$, không cho biết riêng từng chiều dài 195, 196, 197 hay 198 mm như danh sách ban đầu.

Câu 3

a) Tại trạm Seattle–Tacoma, trung bình của ba giá trị lượng giáng thủy được ghi từ ngày 09 đến ngày 11/03/2023 xấp xỉ 1,63 mm mỗi ngày; tổng ba số là 4,9 mm.

b) Nếu bản chép mất ngày 10/03, trung bình của hai số còn biết là 0,4 mm. Kết quả này mô tả hai giá trị ngày 09 và 11/03, không phải trung bình của cả ba ngày.

c) Không đồng ý. Số 0 là một giá trị được ghi nhận; ô mất số liệu không cho biết giá trị của ngày ấy. Điền 0 sẽ tự thêm một thông tin chưa biết và làm sai trung bình của ba ngày trong tình huống này.

Câu 4

Tổng là 30 phút, trung bình 6 phút và trung vị 5 phút. Viết cùng các thời gian ấy bằng giây làm mỗi giá trị nhân với 60, nên tổng thành 1800 giây. Số quan sát vẫn là 5; trung bình bằng $\frac{1800}{5} = 360$ giây, đúng bằng 6 phút. Nhân với số dương 60 giữ thứ tự, nên giá trị giữa thành $5 \cdot 60 = 300$ giây, đúng bằng 5 phút. Thời gian thực tế không đổi; con số và đơn vị phải được đổi cùng nhau.

Đáp án tự kiểm tra — Chọn đúng câu đã làm

Mỗi đáp án đóng riêng. Đừng mở câu chưa tự làm; mở lời giải nhiệm vụ chính không thay cho lần làm độc lập câu tự kiểm.

Đáp án câu tự kiểm 1 — Vị trí và tần số

trung bình $\frac{17}{3} \approx 5{,}67$; trung vị 5,5; $Q_1 = 4$, $Q_3 = 8$; mốt 4.

Trở lại câu tự kiểm 1 (mở trên trang học liệu).

Đáp án câu tự kiểm 2 — Ghép nhóm

20 thuộc [20; 30); có 5 giá trị trước nhóm này. Tổng tích trung điểm với tần số là 80, cỡ mẫu 6; trung bình ghép nhóm xấp xỉ 13,33.

Trở lại câu tự kiểm 2 (mở trên trang học liệu).

Đáp án câu tự kiểm 3 — Ô thiếu

có 2 giá trị dùng được, trung bình bằng 3. Không tính ô thiếu vào mẫu số.

Trở lại câu tự kiểm 3 (mở trên trang học liệu).

Xem đề bài (mở trên trang học liệu).

Lời giải Nhiệm vụ 20 — Đọc dữ liệu khi cách trình bày thay đổi

Câu 1

Dãy đã sắp là:

$$187;\ 187;\ 191;\ 200;$$
$$200;\ 201;\ 205.$$

a) Trung vị là 200 mm ở vị trí thứ tư.

Nửa dưới gồm 187; 187; 191, nên $Q_1 = 187$ mm. Nửa trên gồm 200; 201; 205, nên $Q_3 = 201$ mm.

Hai mốt là 187 và 200 mm: mỗi giá trị xuất hiện hai lần, còn các giá trị 191, 201, 205 mm đều xuất hiện một lần.

b) Hai số đo 200 mm đứng ở vị trí thứ tư và thứ năm. Quan sát thứ tư ở chính giữa nên không đưa vào hai nửa khi tìm Q_1, Q_3. Quan sát thứ năm vẫn thuộc nửa trên.

Hai cá thể có cùng số đo nhưng hai quan sát vẫn chiếm hai vị trí trong mẫu. Không được bỏ mọi số đo bằng trung vị.

c) Có thể viết: “Trong nhóm bảy cá thể Chinstrap được ghi nhận năm 2008 này, trung vị chiều dài cánh bơi là 200 mm.”

Câu ấy không cho biết từng chiều dài. Những mẫu khác nhau vẫn có thể có cùng trung vị.

Câu 2

a) Tần số tích lũy lần lượt là:

$$0;\ 0;\ 3;\ 5;$$
$$10;\ 14;\ 16;\ 18.$$

Tổng các tích trung điểm với tần số là 3585. Do đó:

$$\overline{x}_{\text{nhóm}} = \frac{3585}{18} \approx 199{,}17 \text{ mm}.$$

b) Các thông tin dùng để tính:

Số cần tìm	Mức tần số tích lũy cần đạt	Khoảng, mm	Số quan sát trước khoảng	Tần số của khoảng
Q_1	4,5	[190; 195)	3	2
Trung vị	9	[195; 200)	5	5
Q_3	13,5	[200; 205)	10	4

Các khoảng rộng 5 mm. Với trung vị, mức tích lũy $r = 9$ nằm giữa 5 và 10, nên chọn [195; 200). Trước nhóm đã có 5, còn thiếu 4 trong tần số nhóm 5; theo giả thiết tích lũy tăng đều, lấy $\frac{4}{5}$ độ rộng 5 mm rồi cộng vào 195 mm. Các tứ phân vị được tính tương tự:

$$\begin{aligned} Q_{1,\text{nhóm}} &= 190 + \frac{4{,}5 - 3}{2} \cdot 5 \\ &= 193{,}75 \text{ mm}. \end{aligned}$$

$$\begin{aligned} M_{e,\text{nhóm}} &= 195 + \frac{9 - 5}{5} \cdot 5 \\ &= 199 \text{ mm}. \end{aligned}$$

$$\begin{aligned} Q_{3,\text{nhóm}} &= 200 + \frac{13{,}5 - 10}{4} \cdot 5 \\ &= 204{,}375 \\ &\approx 204{,}38 \text{ mm}. \end{aligned}$$

c) Các khoảng đều rộng 5 mm. Khoảng [195; 200) là khoảng duy nhất có tần số lớn nhất: 5 cá thể. Hai khoảng liền trước và liền sau có tần số 2 và 4.

$$\begin{aligned} M_{o,\text{nhóm}} &= 195 + \frac{5 - 2}{(5 - 2) + (5 - 4)} \cdot 5 \\ &= 198{,}75 \text{ mm}. \end{aligned}$$

d) Không thể khẳng định có cá thể được đo đúng 198,75 mm. Đây là kết quả tính từ tần số các nhóm để ước lượng vị trí của mốt. Bảng không cung cấp từng số đo trong khoảng.

Có thể viết: “Khoảng từ 195 mm đến dưới 200 mm có nhiều cá thể nhất; mốt tính từ bảng nhóm là 198,75 mm.”

Câu 3

a) Tổng là 44959 lượt thuê trong bảy ngày:

$$\overline{x} = \frac{11977 + 32982}{7} \approx 6422{,}71 \text{ lượt/ngày.}$$

b) Hai trung bình nhóm là 5988,5 và 6596,4 lượt/ngày. Trung bình cộng của chúng bằng 6292,45 lượt/ngày.

Cộng hai trung bình rồi chia tổng cho 2 là tính hai nhóm ngang nhau, dù một nhóm có hai ngày và nhóm kia có năm ngày. Trong cả tuần, hai nhóm chiếm $\frac{2}{7}$ và $\frac{5}{7}$ số ngày. Muốn tính đủ cả bảy ngày, ta khôi phục tổng của từng nhóm rồi lấy $\frac{2\cdot 5988{,}5 + 5\cdot 6596{,}4}{7} = \frac{44959}{7} \approx 6422{,}71$ lượt mỗi ngày. Các tích ở tử chính là tổng của từng nhóm.

c) Không đồng ý. Một người có thể tạo ra nhiều lượt thuê, nên tổng lượt thuê không cho biết số người khác nhau.

Có thể sửa thành: “Từ ngày 01 đến ngày 07/04/2012, hệ thống ghi nhận tổng cộng 44959 lượt thuê xe.”

Câu 4

Mức tích lũy một nửa mẫu là $r = \frac{24}{2} = 12$. Trước nhóm $[12; 24)$ đã có $F = 8$, đến hết nhóm có 16, nên chọn nhóm này. Phần cần thêm là $12 - 8 = 4$, chiếm $\frac{4}{8} = \frac{1}{2}$ tần số nhóm. Khi giả sử tần số tích lũy tăng đều từ đầu trái đến đầu phải, cùng tỉ lệ $\frac{1}{2}$ được dùng cho độ rộng $24 - 12 = 12$. Trung vị ước lượng là $12 + \frac{1}{2}\cdot 12 = 18$. Bảng không cho biết vị trí từng giá trị, nên kết quả không chứng minh đúng bốn giá trị gốc trong nhóm nhỏ hơn 18. Tỉ lệ theo độ rộng là hệ quả của giả thiết tăng đều, không phải điều đã được bảng quan sát xác nhận.

Đáp án tự kiểm tra — Chọn đúng câu đã làm

Mỗi đáp án đóng riêng. Đừng mở câu chưa tự làm; mở lời giải nhiệm vụ chính không thay cho lần làm độc lập câu tự kiểm.

Đáp án câu tự kiểm 1 — Mẫu lẻ

$Q_2 = 8$; nửa dưới 3; 5; 5 cho $Q_1 = 5$; nửa trên 9; 12; 15 cho $Q_3 = 12$.

Trở lại câu tự kiểm 1 (mở trên trang học liệu).

Đáp án câu tự kiểm 2 — Bảng nhóm

cỡ mẫu 10. Ba mốc tần số 2,5; 5; 7,5 đều được nội suy trong khoảng $[10; 20)$, có 2 quan sát trước khoảng và tần số 6. Kết quả: $Q_1 \approx 10{,}83$, trung vị 15, $Q_3 \approx 19{,}17$, mốt 15. Để tính mốt, chọn nhóm $[10; 20)$ có tần số 6, lớn hơn tần số 2 của mỗi nhóm liền kề. Hai độ chênh cùng bằng $6 - 2 = 4$, nên kết quả là $10 + \frac{4}{4+4}\cdot 10 = 15$, đúng tại trung điểm nhóm.

Trở lại câu tự kiểm 2 (mở trên trang học liệu).

Đáp án câu tự kiểm 3 — Gộp dữ liệu

tổng 65, cỡ mẫu 5, trung bình 13. Trung bình cộng của hai trung bình nhóm là 12,5; kết quả này không tính đến việc hai nhóm có số giá trị khác nhau.

Trở lại câu tự kiểm 3 (mở trên trang học liệu).

Xem đề bài (mở trên trang học liệu).

Lời giải Nhiệm vụ 21 — Một kết luận có nói đúng điều dữ liệu cho biết không?

Câu 1

a) Số trung bình:

$$\overline{x} = \frac{4754}{24} \approx 198{,}08 \text{ mm}.$$

Hai vị trí giữa, thứ 12 và thứ 13, đều mang giá trị 198 mm, nên trung vị bằng 198 mm.

Có sáu mốt: 187, 193, 198, 201, 202, 203 mm. Mỗi giá trị ấy xuất hiện hai lần; mỗi giá trị khác chỉ xuất hiện một lần. Vì vậy, đây là trường hợp nhiều mốt, không phải mọi giá trị khác nhau đều có cùng tần số.

b) Một cách viết tham khảo: “Với 24 cá thể Chinstrap được ghi nhận năm 2009 ở khu vực quần đảo Palmer, Nam Cực, trung bình cộng của hai chiều dài cánh bơi ở giữa dãy đã sắp là 198 mm. Các số đo trong nhóm nằm từ 187 mm đến 212 mm.”

Người viết chọn trung vị 198 mm để người đọc biết mức chiều dài ở giữa dãy đã sắp. Nêu thêm số nhỏ nhất và số lớn nhất giúp người đọc thấy các cá thể không có cùng một chiều dài, thay vì hình dung cả nhóm đều bằng 198 mm. Tuy vậy, hai câu trên chưa cho biết từng chiều dài xuất hiện bao nhiêu lần hoặc các số đo tập trung ở đâu trong khoảng từ 187 mm đến 212 mm.

Có thể chọn thông tin khác nếu giải thích được nó giúp trả lời câu hỏi nào. Khi viết, cần giữ đúng nhóm chim, năm ghi nhận và đơn vị; không biến một số đại diện thành chiều dài của mọi cá thể.

c) Dùng số trung bình chính xác nhân với 24 sẽ thu lại tổng 4754 mm. Không làm tròn trước khi nhân: nếu dùng 198,08, tích chỉ bằng 4753,92 mm.

d) Dữ liệu chỉ gồm 24 quan sát trong một địa điểm và thời gian xác định, không đủ để nói về mọi chim Chinstrap trên thế giới. Trung vị cũng không nói các cá thể có chiều dài bằng nhau.

Có thể viết: “Trong phần trích 24 cá thể Chinstrap được ghi nhận năm 2009 này, trung vị chiều dài cánh bơi là 198 mm.”

Câu 2

a) Tổng các tích trung điểm với tần số là 4765:

$$\overline{x}_{\text{nhóm}} = \frac{4765}{24} \approx 198{,}54 \text{ mm}.$$

Tần số tích lũy là 3, 8, 14, 20, 22, 24. Mức tần số tích lũy cần đạt trung vị là 12, được xác định trong khoảng $[195; 200)$. Trước khoảng có 8 quan sát; khoảng có 6 quan sát và rộng 5 mm.

$$M_{e,\text{nhóm}} = 195 + \frac{12-8}{6}\cdot 5$$
$$\approx 198{,}33 \text{ mm}.$$

Trung bình mẫu gốc khoảng 198,08 mm; trung vị mẫu gốc bằng 198 mm. Hai kết quả từ bảng nhóm đều khác các kết quả tương ứng của mẫu gốc.

b) Nhận xét sai ở từ "duy nhất". Hai khoảng [195; 200) và [200; 205) đều có 6 cá thể, cùng đạt tần số lớn nhất.

c) Sự khác nhau không tự chứng minh có phép tính sai. Danh sách cho biết chính xác hai số đo ở vị trí giữa. Bảng nhóm chỉ cho biết bao nhiêu số đo thuộc mỗi khoảng, không giữ vị trí chính xác của từng số đo bên trong khoảng. Trung vị tính từ bảng nhóm vì thế là kết quả ước lượng.

Câu 3

a) Tổng sáu giá trị được ghi là 4,6 mm:

$$\overline{x} = \frac{4{,}6}{6} \approx 0{,}77 \text{ mm}.$$

Dãy đã sắp là 0; 0; 0; 0; 1; 3,6, nên trung vị bằng 0 mm.

b) Không đồng ý. Các ngày 19, 21 và 24/03 có số 0 kèm T, tức vẫn ghi nhận vết giáng thủy. Bỏ cột ghi chú đã làm mất thông tin ấy.

Có thể nói bảng có bốn giá trị lượng giáng thủy được ghi bằng 0; không thể từ đó đổi thành bốn ngày hoàn toàn không có giáng thủy.

c) Có thể viết: "Tại trạm Seattle–Tacoma, trung bình của sáu giá trị lượng giáng thủy được ghi từ ngày 19 đến ngày 24/03/2023 xấp xỉ 0,77 mm mỗi ngày."

Kết quả được tính từ các số đã ghi. Những lượng rất nhỏ được ghi bằng 0 kèm T không được khôi phục thành lượng nước cụ thể trong phép tính. Vì vậy, không gọi kết quả này là trung bình thực tế chính xác tuyệt đối.

Câu 4

Tổng bằng 48, nên trung bình bằng 6. Hai giá trị ở giữa là 2 và 10, nên trung vị cũng bằng 6. Không có quan sát nào giữa 2 và 10. Vì thế đổi sang trung vị vẫn không cho thấy hai cụm. Có thể viết: "Bốn quan sát nằm ở 1–2, bốn quan sát ở 10–11; trung bình và trung vị cùng bằng 6, ở khoảng trống giữa hai cụm." Đối chứng này bác bỏ từ "cứ" trong phát biểu; lựa chọn phụ thuộc câu hỏi và cách dữ liệu phân bố.

Đáp án tự kiểm tra — Chọn đúng câu đã làm

Mỗi đáp án đóng riêng. Đừng mở câu chưa tự làm; mở lời giải nhiệm vụ chính không thay cho lần làm độc lập câu tự kiểm.

Đáp án câu tự kiểm 1 — Chọn số đại diện

trung bình 4,5, trung vị 4. Bốn giá trị khác nhau là 1, 3, 5, 9 đều xuất hiện hai lần; theo quy ước đã học, mẫu không có mốt. Việc các số đều lặp lại không đủ để kết luận mẫu có mốt. Trung vị 4 không xuất hiện trong mẫu.

Trở lại câu tự kiểm 1 (mở trên trang học liệu).

Đáp án câu tự kiểm 2 — Bảng nhóm

cỡ mẫu 12, tổng các tích trung điểm với tần số là 240, nên trung bình ghép nhóm bằng 20. Để tìm trung vị, mức tích lũy cần đạt là $r = 6$. Nhóm $[10; 20)$ có $F = 2$, $f = 4$, nên mức 6 đạt ở cuối nhóm. Theo nội suy tăng đều, trung vị ước lượng là $10 + \frac{6-2}{4} \cdot 10 = 20$. Hai khoảng đông nhất là $[10; 20)$ và $[20; 30)$. Chưa biết từng giá trị trong các khoảng nên chưa xác định chính xác trung bình mẫu gốc.

Trở lại câu tự kiểm 2 (mở trên trang học liệu).

Đáp án câu tự kiểm 3 — Ghi chú số đo

trung bình của ba số được ghi bằng 0,8 mm. Ngày có 0 kèm T vẫn ghi nhận vết giáng thủy, nên không thể gọi cả hai ngày ghi 0 là hoàn toàn không có giáng thủy.

Trở lại câu tự kiểm 3 (mở trên trang học liệu).

Xem đề bài (mở trên trang học liệu).

Ôn nền

Lời giải Nhiệm vụ 22 — Tự kiểm tra phép tính cơ bản

a) Tổng $4 + 6 + 8 + 10 = 28$, trung bình $\frac{28}{4} = 7$. Ta chia cho bốn vì đang dùng đủ bốn giá trị của bảng.

b) Hai giá trị lớn hơn 7 là 8 và 10, chiếm $\frac{2}{4} = 50\%$. Như vậy, một nửa số giá trị trong bảng lớn hơn ngưỡng đã cho.

c) Còn ba giá trị 4; 6; 8, trung bình $\frac{18}{3} = 6$. Đây là trung bình của ba giá trị đã ghi. Chưa biết giá trị ở ô cuối nên chưa thể tính trung bình của đủ bốn ô; không coi ô trống là một giá trị bằng 0.

Xem đề bài (mở trên trang học liệu).

Lời giải Nhiệm vụ 23 — Tự kiểm tra cách đọc khoảng

a) Dãy đã sắp: 9; 10; 14; 15; 15. Dãy vẫn có năm giá trị, gồm cả hai lần xuất hiện của số 15.

b) Xếp từng giá trị: 9 thuộc $[5; 10)$; 10 và 14 thuộc $[10; 15)$; cả hai giá trị 15 thuộc $[15; 20)$. Vì thế tần số lần lượt là 1; 2; 2, có tổng 5, đúng bằng cỡ mẫu. Kiểm tổng giúp phát hiện một số sai sót, nhưng vẫn phải kiểm từng giá trị thuộc đúng khoảng: tổng đúng chưa loại trừ khả năng bỏ sót một giá trị và đếm lặp giá trị khác.

Xem đề bài (mở trên trang học liệu).

Lời giải Nhiệm vụ 24 — Tự tính trung bình của hai nhóm

a) Tổng thời gian chờ là $4 \cdot 5 + 6 \cdot 10 = 80$ phút. Chia cho tổng $4 + 6 = 10$ khách được trung bình 8 phút. Tử số cộng thời gian của đủ mười khách, mẫu số đếm chính mười khách ấy.

b) Trung bình cộng của hai trung bình là 7,5 phút, không bằng 8 phút. Nếu mỗi nhóm cùng có $n > 0$ khách và trung bình lần lượt a, b, trung bình chung là $\frac{na+nb}{2n} = \frac{a+b}{2}$. Vì vậy cách chia đôi đúng khi hai nhóm có cùng cỡ. Với cỡ khác nhau nhưng hai trung bình bằng nhau, hai cách cũng trùng kết quả; không suy ngược "chia đôi ra đúng" thành "hai nhóm cùng cỡ".

Xem đề bài (mở trên trang học liệu).

Lời giải Nhiệm vụ 25 — Tự chọn nhóm và nội suy

a) Tần số tích lũy là 3; 13; 20. Đối với trung vị, $r = 10$ nằm sau 3 và trước 13, nên chọn $[8; 16)$: $F = 3$, $f = 10$, độ rộng 8 phút. Giả sử tần số tích lũy tăng đều trong nhóm, ta có

$$M_e = 8 + \frac{10 - 3}{10} \cdot 8 = 13{,}6 \text{ phút}.$$

Đối với Q_3, $r = \frac{3 \cdot 20}{4} = 15$ nằm sau 13 và trước 20, nên phải đổi sang nhóm $[16; 24)$: $F = 13$, $f = 7$, độ rộng 8 phút. Với cùng giả thiết nội suy,

$$Q_3 = 16 + \frac{15 - 13}{7} \cdot 8 = \frac{128}{7} \approx 18{,}29 \text{ phút}.$$

b) Không thể khẳng định. Bảng giữ số khách trong mỗi khoảng, không giữ từng thời gian. Nếu ba khách chờ 1 phút, mười khách chờ 9 phút và bảy khách chờ 17 phút, bảng vẫn như đề nhưng hai giá trị ở giữa đều bằng 9 phút. Trung vị mẫu gốc khi đó là 9 phút, khác 13,6 phút. Kết quả nội suy phụ thuộc giả thiết, không khôi phục chính xác mẫu gốc.

Xem đề bài (mở trên trang học liệu).

Sửa lỗi — Đối chiếu đúng ý đã tự làm

Lời giải Nhiệm vụ 26 — Tự làm với mẫu Gentoo nhỏ

a) $n = 12$. Dãy đã sắp:

$$208;\ 208;\ 208;\ 210;\ 213;\ 216;$$
$$217;\ 219;\ 220;\ 220;\ 222;\ 225.$$

b) Tổng 2586 mm nên $\overline{x} = 215{,}5$ mm. Gọi x_i là số đo ở vị trí thứ i trong dãy đã sắp. Hai vị trí giữa của toàn mẫu là 6, 7; mỗi nửa có sáu giá trị:

$$Q_1 = \frac{x_3 + x_4}{2} = \frac{208 + 210}{2} = 209;$$
$$M_e = Q_2 = \frac{x_6 + x_7}{2} = \frac{216 + 217}{2} = 216{,}5;$$
$$Q_3 = \frac{x_9 + x_{10}}{2} = \frac{220 + 220}{2} = 220.$$

Các kết quả mô tả chiều dài cánh bơi của mười hai cá thể Gentoo đang xét, có đơn vị mm. Mốt là 208 mm: giá trị này xuất hiện ba lần, 220 mm xuất hiện hai lần, còn mỗi giá trị khác xuất hiện một lần.

c) Chỉ dùng 11 số đo đầu, không đưa số đo cuối 217 mm của danh sách ban đầu vào mẫu này. Sau khi sắp xếp mẫu 11 giá trị, trung vị là giá trị ở vị trí thứ 6, bằng 216 mm. Để tìm hai tứ phân vị còn lại bằng trung vị của hai nửa, nửa dưới gồm năm số đứng trước vị trí thứ 6, nửa trên gồm năm số đứng sau. Quan sát thứ 6 vẫn thuộc mẫu và đã được dùng để tìm Q_2; trung vị của nửa dưới là 208, của nửa trên là 220. Vậy $Q_1 = 208$, $Q_2 = 216$, $Q_3 = 220$ mm.

d) Mỗi mẫu có tổng 18, trung bình 3 và trung vị 3. A có nửa dưới 0; 2; 2, nửa trên 4; 4; 6, nên $Q_1 = 2$, $Q_3 = 4$. B có nửa dưới 0; 0; 3, nửa trên 3; 6; 6, nên $Q_1 = 0$, $Q_3 = 6$. Các giá trị trong mỗi

nửa khác nhau; hai tứ phân vị phản ánh sự khác nhau ấy dù trung bình, trung vị và hai giá trị đầu, cuối bằng nhau.

Xem đề bài (mở trên trang học liệu).

Lời giải Nhiệm vụ 27 — Tự kiểm tra ghép nhóm và nội suy

a) Với riêng 12 giá trị, tần số theo bảy khoảng từ $[200;205)$ đến $[230;235)$ là $0;1;4;2;3;2;0$. Tổng bằng 12. Hai mốt gốc là 210 và 220 mm: mỗi giá trị xuất hiện ba lần, 225 mm xuất hiện hai lần, còn mỗi giá trị khác xuất hiện một lần.

b) Trung điểm của $[215;220)$ là 217,5 mm. Bảng chỉ cho biết bảy cá thể có chiều dài cánh bơi thuộc khoảng ấy, không cho biết từng số đo; vì vậy không thể nói cả bảy số đo đều bằng trung điểm.

c) Với toàn bộ 34 giá trị, tần số tích lũy là $0;4;10;17;28;31;34$. Tổng các tích tần số với trung điểm là 7455:

$$\overline{x}_{\text{nhóm}} = \frac{7455}{34} \approx 219{,}26 \text{ mm}.$$

$$\begin{aligned} Q_{1,\text{nhóm}} &= 210 + \frac{8{,}5-4}{6}\cdot 5 \\ &= 213{,}75 \text{ mm}; \\ M_{e,\text{nhóm}} &= 215 + \frac{17-10}{7}\cdot 5 \\ &= 220 \text{ mm}; \\ Q_{3,\text{nhóm}} &= 220 + \frac{25{,}5-17}{11}\cdot 5 \\ &\approx 223{,}86 \text{ mm}. \end{aligned}$$

Mức tần số tích lũy cần đạt cho trung vị là $r = \frac{34}{2} = 17$, đúng bằng mức đến hết nhóm $[215;220)$. Với nhóm này, phần tần số cần thêm là $17-10=7$, bằng toàn bộ tần số nhóm, nên lấy đủ độ rộng 5 mm và được 220 mm. Nhóm kế tiếp $[220;225)$ có $F=17$, $f=11$, cho $220+\frac{17-17}{11}\cdot 5 = 220$ mm. Hai nhóm kề đều có tần số dương nên hai cách cho cùng biên; không đếm một quan sát vào hai nhóm.

d) Các nhóm rộng đều 5 mm; nhóm $[220;225)$ có tần số lớn nhất duy nhất 11:

$$\begin{aligned} M_{o,\text{nhóm}} &= 220 + \frac{11-7}{(11-7)+(11-3)}\cdot 5 \\ &\approx 221{,}67 \text{ mm}. \end{aligned}$$

Kết quả này mô tả mẫu ghép nhóm của 34 số đo cánh bơi, được nội suy từ bảng; không chứng minh có một quan sát mang đúng giá trị ấy.

Xem đề bài (mở trên trang học liệu).

Lời giải Nhiệm vụ 28 — Đếm giá trị, tính trung bình và giới hạn kết luận

a) Có 44 dòng, với 43 số đo chiều dài dùng được và một ô thiếu. Dòng có ô thiếu vẫn ghi một cá thể, nhưng không cung cấp số đo cánh bơi để đưa vào phép tính. Ô thiếu không phải một phép đo bằng 0.

b) Trung bình hai phần lần lượt là $\frac{4375}{20} = 218{,}75$ mm và $\frac{5017}{23} \approx 218{,}13$ mm. Trung bình chung:

$$\begin{aligned}\overline{x} &= \frac{4375 + 5017}{20 + 23} \\ &= \frac{9392}{43} \\ &\approx 218{,}42 \text{ mm}.\end{aligned}$$

Dùng hai trung bình chưa làm tròn, nhân lần lượt với 20 và 23 sẽ khôi phục tổng 4375 và 5017 mm. Cộng hai tổng rồi chia cho 43 số đo cho trung bình chung. Không dùng 24 cho phần sau vì có một ô thiếu. Cũng không chia đôi tổng hai trung bình: hai phần chiếm $\frac{20}{43}$ và $\frac{23}{43}$ số đo, không phải mỗi phần một nửa.

c) "Trong 43 giá trị hợp lệ của phần trích Gentoo 2009, chiều dài cánh bơi trung bình khoảng 218,42 mm." Trung bình không làm mọi quan sát bằng nhau và không tự mở rộng phạm vi ra mọi cá thể Gentoo.

d) Bốn giá trị 212, 221, 224, 228 mm khác nhau và đều xuất hiện một lần. Các tần số đều bằng 1, nên mẫu không có mốt theo quy ước đã học.

e) Mẫu giả định ban đầu có tổng 35, trung bình 7 phút và trung vị 6 phút. Sau khi tăng giá trị cuối thêm 25, tổng là 60, trung bình 12 phút, trung vị vẫn 6 phút. Câu hỏi về người ở giữa dùng trung vị; câu hỏi về tổng dùng tổng hoặc tích cỡ mẫu với trung bình chưa làm tròn. Giá trị 36 là một quan sát đã được đề cho; muốn loại phải có căn cứ về sai dữ liệu hoặc thay đổi phạm vi, không phải vì nó làm mức chia đều tăng.

Xem đề bài (mở trên trang học liệu).

Lời giải Nhiệm vụ 29 — Đọc ký hiệu T

a) Tổng các giá trị đã ghi là 2,3 mm. Trung bình $\frac{2{,}3}{3} \approx 0{,}77$ mm/ngày. Dãy sắp là 0; 0; 2,3, nên trung vị bằng 0 mm.

b) Không. Ngày 10/02 có ký hiệu T, cho biết vết giáng thủy. Có thể viết: "Trong ba ngày tại trạm Seattle–Tacoma đang xét, hai ngày có lượng giáng thủy được ghi bằng 0; riêng ngày 10/02 được gắn ký hiệu vết giáng thủy." Không xóa ký hiệu khi diễn giải.

c) Trung vị bằng 0 vì giá trị ở giữa dãy 0; 0; 2,3 là 0. Tổng vẫn bằng 2,3 mm vì phép cộng tính cả lượng giáng thủy dương của ngày 09/02. Muốn mô tả cả năm, cần dữ liệu của các ngày còn lại; ba ngày này chưa đủ để kết luận cho cả năm.

Xem đề bài (mở trên trang học liệu).

Nguồn và cách dùng dữ liệu

Mục này cho biết kiến thức toán dựa vào tài liệu nào, số liệu đến từ đâu và ZO Math đã chọn, trình bày số liệu ra sao. Em có thể dùng các liên kết để tra cách ghi số đo hoặc đọc thêm; không cần học thuộc tên tác giả, mã cột hay mã trạm.

Tài liệu toán học

- **Toán 10, Tập một — Kết nối tri thức với cuộc sống**, Nhà xuất bản Giáo dục Việt Nam, Bài 13, trang 78–81: trung bình, trung vị, tứ phân vị và mốt của mẫu chưa ghép nhóm.
- **Toán 11, Tập một — Kết nối tri thức với cuộc sống**, Nhà xuất bản Giáo dục Việt Nam, Bài 8, trang 58–60; Bài 9, trang 62–66: mẫu ghép nhóm, các số đặc trưng và điều kiện áp dụng.
- **Chương trình giáo dục phổ thông môn Toán 2018**, Bộ Giáo dục và Đào tạo, trang 85 và 101–102: yêu cầu tính, giải thích ý nghĩa và sử dụng các số đặc trưng của mẫu không ghép nhóm và ghép nhóm, phần lớp 10 và lớp 11.
- **Toán 10 — Sách giáo viên, Kết nối tri thức với cuộc sống**, Nhà xuất bản Giáo dục Việt Nam, trang 142–145: ý nghĩa, cách lựa chọn các số đặc trưng, hướng dẫn dạy mốt và lời giải liên quan.
- **Toán 11 — Sách giáo viên, Kết nối tri thức với cuộc sống**, Nhà xuất bản Giáo dục Việt Nam, trang 103–105: cách dẫn dắt về mốt ghép nhóm, giải thích kết quả và phân biệt nhóm chứa mốt với giá trị tính được.
- **Bài tập Toán 10, Tập một — Kết nối tri thức với cuộc sống**, Nhà xuất bản Giáo dục Việt Nam, trang 75, 77, 126: phần nhắc kiến thức về mốt, bài 5.10 và lời giải.
- **Bài tập Toán 11, Tập một — Kết nối tri thức với cuộc sống**, Nhà xuất bản Giáo dục Việt Nam, trang 48, 50–52, 128–129: quy tắc mốt ghép nhóm, các bài 3.8, 3.15, 3.20–3.21, 3.25 và lời giải.

Chương trình xác định yêu cầu cần đạt; sách giáo khoa cung cấp căn cứ về định nghĩa, quy tắc và công thức; sách giáo viên cung cấp gợi ý về ý nghĩa và cách dẫn dắt. Các bộ dữ liệu dưới đây cung cấp số đo, đơn vị và phạm vi quan sát.

Lời giảng, nhiệm vụ, lời giải và hình minh họa do ZO Math biên soạn; không sao chép bài tập hoặc hình sách.

Lưu ý về định nghĩa và quy ước mốt. Toán 10, Tập một, bộ Kết nối tri thức với cuộc sống, trang 81, định nghĩa mốt theo tần số lớn nhất và nêu quy ước không có mốt khi các giá trị xuất hiện với tần số như nhau. Với mẫu có nhiều giá trị khác nhau, cần đọc cả định nghĩa lẫn quy ước này: việc các giá trị đều lặp lại chưa đủ để kết luận có mốt.

Nếu mẫu chỉ có một giá trị khác nhau, cần phân biệt điều đếm được với cách gọi tên: chẳng hạn, trong 5; 5; 5, chắc chắn giá trị 5 xuất hiện ba lần. Việc gọi đó là "mốt" hay áp dụng lời quy ước "các tần số như nhau" cần được làm rõ trong định nghĩa mà để sử dụng; không suy ra một quy tắc mới chỉ từ cách đếm. Trang sách nêu trên không giải thích riêng trường hợp này.

Với mẫu ghép nhóm, công thức ở Toán 11, Tập một, trang 66, áp dụng cho các nhóm cùng độ rộng; sách cũng nêu quy ước không có mốt khi tất cả tần số nhóm bằng nhau. Khi có nhiều nhóm cùng đạt tần số lớn nhất, xét từng nhóm cùng tần số của hai nhóm liền kề: nếu mẫu số khác 0 thì tính và đối chiếu các kết quả; nếu mẫu số bằng 0 thì ghi rõ công thức không xác định ở nhóm đó. Hai nhóm kề có thể cho cùng một giá trị tại biên, như ví dụ trong phần mốt ghép nhóm (mở trên trang học liệu). Không suy số giá trị mốt từ số nhóm có tần số lớn nhất và không tự dùng công thức cho các nhóm khác độ rộng.

Palmer Penguins

Đây là bộ dữ liệu về chim cánh cụt được nghiên cứu ở khu vực quần đảo Palmer, Nam Cực. Adélie, Chinstrap và Gentoo là tên ba loài có trong bảng. Các bài tập sử dụng chiều dài cánh bơi — phần cánh có dạng mái chèo giúp chim bơi — đo bằng mm; trong bảng gốc, cột này mang tên flipper_length_mm. Mỗi dòng ghi thông tin một cá thể.

Trang dữ liệu và mô tả biến · Tệp bảng số liệu CSV · CC0 1.0.

Ghi công Allison Horst, Alison Hill, Kristen Gorman. Dữ liệu gốc do Kristen Gorman và chương trình nghiên cứu sinh thái dài hạn tại trạm Palmer (Palmer Station LTER) cung cấp; trang nguồn dẫn tới các bộ dữ liệu lưu tại Environmental Data Initiative (EDI). Nghiên cứu liên quan: Gorman, Williams và Fraser (2014), Ecological Sexual Dimorphism and Environmental Variability within a Community of Antarctic Penguins.

ZO Math chọn các số đo cánh bơi theo loài và năm nghiên cứu để lập bài học. Ô thiếu vẫn được giữ trong hồ sơ; khi tính với chiều dài cánh bơi, chỉ dùng các ô có số đo phù hợp, không thay ô thiếu bằng 0. Các bảng ghép nhóm được lập từ những số đo ấy; kết quả gần đúng chỉ làm tròn ở bước cuối. Những nhóm nhỏ dùng trong bài được chọn để học thao tác, không phải mẫu lấy ngẫu nhiên để đại diện cho toàn bộ loài.

Bike Sharing

Bộ dữ liệu này ghi số lượt thuê xe đạp của hệ thống Capital Bikeshare tại Hoa Kỳ trong các năm 2011–2012. Một lượt thuê là một lần sử dụng dịch vụ, không phải một người dùng khác nhau. UCI Machine Learning Repository là kho lưu trữ và giới thiệu bộ dữ liệu, được liên kết dưới đây.

Fanaee-T, H. và Gama, J., Bike Sharing — UCI Machine Learning Repository, DOI dữ liệu, CC BY 4.0.

ZO Math dùng tệp day.csv, trong đó mỗi dòng ứng với một ngày, chọn các cột ngày, số lượt thuê và thông tin cho biết ngày đó có phải ngày làm việc hay không; trích theo tháng, tính tổng và các số đặc trưng, dịch nhãn sang tiếng Việt. Không thay các giá trị quan sát. Việc ghi công không hàm ý tác giả hoặc UCI bảo trợ học liệu.

Dữ liệu mô tả một hệ thống thuê xe và giai đoạn cụ thể. Không dùng lượt thuê để suy trực tiếp số người khác nhau, năng lực phục vụ cần có ở mọi thời điểm hoặc nguyên nhân của sự khác biệt.

NOAA/NCEI

NOAA là Cơ quan Quản lý Khí quyển và Đại dương Quốc gia Hoa Kỳ; NCEI là Trung tâm Thông tin Môi trường Quốc gia thuộc NOAA. Bộ GHCN-Daily tập hợp số liệu khí tượng theo ngày. Các số liệu được sử dụng ở đây ghi lượng giáng thủy tại trạm sân bay Seattle–Tacoma, bang Washington, Hoa Kỳ; USW00024233 là mã giúp tìm đúng trạm trong nguồn.

Menne và cộng sự, GHCN-Daily, Version 3, giải thích các cột và ký hiệu trong dữ liệu, thông tin trạm USW00024233.

PRCP là mã cột lượng giáng thủy: nước từ mưa, tuyết… được quy đổi thành lượng nước tương đương. Các số liệu được chọn thuộc năm 2023, được lấy theo hệ mét và trình bày bằng mm. Ký hiệu T cho biết vết giáng thủy — một lượng rất nhỏ chưa thể hiện bằng số đo dương ở độ chính xác đang dùng.

ZO Math giữ các ký hiệu đi kèm số đo. Tổng và trung bình được tính từ những con số đã ghi; không tự đặt một lượng nước cụ thể thay cho T. Khi diễn giải kết quả, vẫn phải đọc ký hiệu này, không coi mọi ngày ghi 0 là hoàn toàn không có giáng thủy.

Chính sách dữ liệu mở của NCEI phân biệt dữ liệu chính phủ Hoa Kỳ thuộc phạm vi công cộng với dữ liệu bên thứ ba có thể có điều kiện riêng. Không suy rằng mọi dữ liệu trên NCEI đều mang CC0. Học liệu chỉ sử dụng phần dữ liệu trạm Hoa Kỳ đã nêu, không sử dụng logo hoặc ảnh của cơ quan.

Các nguồn dữ liệu được truy cập ngày 22/10/2026. Khi diễn giải số liệu, cần giữ đúng đơn vị, mẫu và thời điểm quan sát.

HỌC TIẾP CÙNG ZO MATH

Chọn điều em muốn học tiếp.

Em có thể bắt đầu với một nội dung đang học, muốn ôn lại hoặc còn vướng khi tự giải bài.

Mở học liệu của khóa học

Bảo trợ ZO Math

Nếu học liệu hữu ích với bạn, bạn có thể bảo trợ ZO Math tiếp tục biên soạn và chia sẻ.

Ngân hàng
Vietcombank

Số tài khoản
0601000137768

Chủ tài khoản
Nguyễn Tấn Nhựt

Nội dung chuyển khoản
BAO TRO ZO MATH

Xem thông tin bảo trợ ZO Math

Phiên bản nội dung 0.4 · Ngày tạo PDF: 04/11/2026

© 2026 ZO Math. Bản quyền nội dung thuộc ZO Math.