

Chương 5

THỐNG KÊ

§1. BẢNG PHÂN BỐ TẦN SỐ VÀ TẦN SUẤT

I. Tóm tắt lí thuyết

1. Bảng phân bố tần số và tần suất


Giả sử dãy n số liệu thống kê đã cho có k giá trị khác nhau ($k \leq n$). Gọi x_i là một giá trị bất kì trong k giá trị đó, ta có:

- Số lần xuất hiện giá trị x_i trong dãy số liệu đã cho được gọi là **tần số** của giá trị đó, kí hiệu là n_i .
- Số $f_i = \frac{n_i}{n}$ được gọi là **tần suất** của giá trị x_i .

2. Bảng phân bố tần số và tần suất ghép lớp

Giả sử dãy n số liệu thống kê đã cho được phân vào k lớp ($k < n$). Xét lớp thứ i ($i = 1, 2, \dots, k$) trong k lớp đó, ta có:

- Số n_i các số liệu thống kê thuộc lớp thứ i được gọi là **tần số của lớp** đó.
- Số $f_i = \frac{n_i}{n}$ được gọi là **tần suất của lớp** thứ i .

 Trong các bảng phân bố tần suất, tần suất được tính ở dạng tỉ số phần trăm.

II. Các dạng toán

Dạng 1. Bảng phân bố tần số và tần suất

Bảng phân bố tần số gồm hai dòng (hoặc hai cột). Dòng (cột) đầu ghi các giá trị khác nhau của mẫu số liệu. Dòng (cột) thứ hai ghi tần số (số lần xuất hiện của mỗi giá trị trong các số liệu thống kê) tương ứng. Nếu bổ sung dòng (cột) thứ ba ghi tần suất (tỉ số % giữa tần số và tổng số liệu thống kê) thì ta được bảng phân bố tần số và tần suất.

Để lập bảng phân bố tần số và tần suất từ bảng số liệu thống kê ban đầu, ta thực hiện các bước sau:

- Sắp thứ tự các giá trị trong các số liệu thống kê;
- Tính tần số n_i của các giá trị x_i bằng cách đếm số lần x_i xuất hiện;
- Tính tần suất f_i của x_i theo công thức $f_i = \frac{n_i}{n}$;
- Đặt các số liệu x_i , n_i , f_i vào bảng ta thu được bảng phân bố tần số và tần suất.

Ví dụ 1. Cho số liệu thống kê ghi trong bảng sau

Thời gian hoàn thành một sản phẩm ở một nhóm công nhân (đơn vị: phút)

42	42	42	42	44	44	44	44	44	45
45	45	45	45	45	45	45	45	45	45
45	45	45	45	45	45	45	45	45	54
54	54	50	50	50	50	48	48	48	48
48	48	48	48	48	48	50	50	50	50

- Hãy lập bảng phân bố tần số và tần suất của bảng trên.
- Trong 50 công nhân được khảo sát, những công nhân có thời gian hoàn thành một sản phẩm từ 45 phút đến 50 phút chiếm bao nhiêu phần trăm.

Lời giải.

- Bảng phân bố tần số và tần suất

Thời gian hoàn thành một sản phẩm ở một nhóm công nhân

Thời gian (phút)	42	44	45	48	50	54	Cộng
Tần số	4	5	20	10	8	3	50
Tần suất (%)	8	10	40	20	16	6	100%

- Từ bảng phân bố tần số và tần suất ở trên, suy ra trong 50 công nhân được khảo sát, những công nhân có thời gian hoàn thành một sản phẩm từ 45 phút đến 50 phút chiếm 76%.

Ví dụ 2. Khi điều tra về năng suất của một giống lúa mới, điều tra viên ghi lại năng suất (tạ / ha) của giống lúa đó trên 40 thửa ruộng có cùng diện tích 1 ha trong bảng sau:

30	32	32	34	38	36	38	36
40	30	40	40	34	38	36	36
38	40	30	40	32	30	30	30
40	38	38	34	34	32	32	36
34	40	34	30	38	38	32	32

- a) Hãy lập bảng phân bố tần số và tần suất của bảng trên.
- b) Trong 40 thửa ruộng được khảo sát, hãy cho biết những thửa ruộng có năng suất cao nhất chiếm bao nhiêu phần trăm.

Lời giải.

- a) Bảng phân bố tần số và tần suất

Năng suất của giống lúa mới của 40 thửa ruộng

Năng suất (tạ / ha)	30	32	34	36	38	40	Cộng
Tần số	7	7	6	5	8	7	40
Tần suất (%)	17,5	17,5	15,0	12,5	20,0	17,5	100%

- b) Từ bảng phân bố tần số và tần suất ở trên, suy ra trong 40 thửa ruộng được khảo sát, những thửa ruộng có năng suất cao nhất chiếm 17,5%.

BÀI TẬP TỰ LUYỆN

Bài 1. Điều tra về tuổi nghề của 30 công nhân được chọn ra từ 150 công nhân của một nhà máy A. Người ta thu được bảng số liệu ban đầu như sau:

7	2	5	9	7	4	3	8	10	4
2	4	4	5	6	7	7	5	4	1
9	4	14	2	8	5	5	7	3	8

Hãy lập bảng phân bố tần số và tần suất của bảng trên.

Lời giải. Bảng phân bố tần số và tần suất

Tuổi nghề của 30 công nhân của một nhà máy A

Tuổi nghề (năm)	1	2	3	4	5	6	7	8	9	10	14	Cộng
Tần số	1	3	2	6	5	1	5	3	2	1	1	30
Tần suất (%)	3,3	10	6,7	20	16,7	3,3	16,7	10	6,7	3,3	3,3	100%

Bài 2. Cho bảng số liệu thống kê năng suất lúa hè thu (tạ / ha) của 30 tỉnh như sau:

25	30	25	30	35	35	40	40	45	25
30	30	40	25	45	45	35	25	35	40
35	35	40	40	30	35	35	35	40	30

- a) Hãy lập bảng phân bố tần số và tần suất của bảng trên.
- b) Nhận xét về xu hướng tập trung của các số liệu thống kê.

Lời giải.

a) Bảng phân bố tần số và tần suất

Năng suất lúa hè thu của 30 tỉnh

Năng suất (tạ / ha)	Tần số	Tần suất (%)
25	5	16,6
30	6	20
35	9	30
40	7	23,4
45	3	10
Cộng	30	100%

b) Từ bảng phân bố tần số và tần suất ở trên, ta thấy năng suất 35 tạ / ha có tần suất cao nhất 30% nên ta nói số liệu thống kê có xu hướng tập trung vào 35 tạ / ha.

Bài 3. Thống kê số con trong mỗi gia đình của 60 gia đình trong một quận được cho ở bảng sau:

2	1	4	2	3	0	2	3	4	2
2	5	1	2	2	3	3	5	7	2
3	4	4	2	1	2	3	2	2	4
6	5	3	4	4	7	2	1	1	5
6	3	5	2	2	3	4	3	5	4
3	3	5	7	2	1	3	4	6	2

a) Hãy lập bảng phân bố tần số và tần suất của bảng trên.

b) Nhận xét về xu hướng tập trung của các số liệu thống kê.

Lời giải.

a) Bảng phân bố tần số và tần suất

Số con trong 60 gia đình của một quận

Số con	Tần số	Tần suất (%)
0	1	1,6
1	6	10
2	17	28,3
3	13	21,7
4	10	16,7
5	7	11,7
6	3	5
7	3	5
Cộng	60	100%

b) Từ bảng phân bố tần số và tần suất ở trên, ta thấy số con là 2 hoặc 3 có tần suất cao nên ta nói số liệu thống kê có xu hướng tập trung vào 2 hoặc 3 con.

Dạng 2. Lập bảng phân bố tần số và tần suất ghép lớp

- Tần số của giá trị x_i (hay một lớp nào đó) là số lần xuất hiện n_i của x_i .
- Tần suất của giá trị x_i (hay một lớp nào đó) là tỉ số $\frac{x_i}{\sum x_i}$.

Ví dụ 3. Nhiệt độ trung bình (đơn vị: °C) của tháng 10 ở địa phương D từ năm 1971 đến 2000 được cho ở bảng sau

27,1	26,9	28,5	27,4	29,1	27,0	27,1	27,4	28,0	28,6
28,1	27,4	27,4	26,5	27,8	28,2	27,6	28,7	27,3	26,8
26,8	26,7	29,0	28,4	28,3	27,4	27,0	27,0	28,3	25,9

Hãy lập bảng phân bố tần số và tần suất ghép lớp của bảng số liệu đã cho?

Lời giải. Bảng phân bố tần số và tần suất ghép lớp được tính như trong bảng sau:

Lớp	Tần số	Tần suất (%)
[25; 27)	6	20
[27; 29)	22	73.33
[29; 31]	2	6.67
Cộng	$n = 30$	100

Ví dụ 4. Kết quả điểm thi môn Toán của 2 lớp 10A1 và 10A2 được cho bởi bảng số liệu sau

Lớp 10A1							
Điểm thi	5	6	7	8	9	10	Cộng
Tần số	5	4	2	26	4	4	45
Lớp 10A2							
Điểm thi	5	6	7	8	9	10	Cộng
Tần số	7	5	6	15	4	1	47

Hãy lập bảng phân bố tần suất kết quả điểm thi môn Toán của hai lớp 10A1 và 10A2. Tìm lớp có điểm 7 chiếm tỉ lệ hơn 50%.

Lời giải. Tần suất kết quả điểm thi môn Toán của hai lớp 10A1 và 10A2 được tính theo như bảng dưới đây:

Lớp 10A1							
Điểm thi	5	6	7	8	9	10	Cộng
Tần số	5	4	2	26	4	4	45
Tần suất (%)	11,11	8,89	4,44	57,78	8,89	8,89	100
Lớp 10A2							
Điểm thi	5	6	7	8	9	10	Cộng
Tần số	7	5	6	15	4	1	47
Tần suất (%)	14,89	10,64	12,77	31,91	8,51	2,12	100

Dựa vào bảng phân bố tần suất đã lập ở trên, ta thấy không có lớp nào có điểm 7 chiếm tỉ lệ hơn 50% .

BÀI TẬP TỰ LUYỆN

Bài 4. Trong một kì thi học sinh giỏi toán (thang điểm là 20), kết quả được cho ở bảng sau:

Điểm	9	10	11	12	13	14	15	16	17	18	19
Tần số	1	1	3	5	8	13	19	24	14	10	2

Lập bảng phân bố tần suất dựa vào số liệu ở bảng trên. Hỏi có bao nhiêu phần trăm thí sinh được chọn vào vòng trong biết rằng điều kiện để vào vòng trong là phải trên 16 điểm?

Lời giải.

Điểm	9	10	11	12	13	14	15	16	17	18	19
Tần số	1	1	3	5	8	13	19	24	14	10	2
Tần suất (%)	1	1	3	5	8	13	19	24	14	10	2

Điều kiện để vào vòng trong là phải trên 16 điểm, nên thí sinh có số điểm là 17, 18, 19 sẽ được vào vòng trong. Vậy số phần trăm thí sinh được vào vòng trong là $14 + 10 + 2 = 26\%$

Bài 5. Trong sổ theo dõi bán hàng ở một cửa hàng bán xe máy có bảng sau:

Số xe bán trong ngày	0	1	2	3	4	5
Tần số	2	13	15	12	7	3

Biết mỗi chiếc xe bán được cửa hàng có lãi 5 triệu đồng, mỗi ngày cửa hàng mất 2 triệu đồng chi phí thuê nhân viên và cơ sở vật chất. Hỏi trong khoảng thời gian đó cửa hàng lãi (hay lỗ) bao nhiêu?

Lời giải. Số tiền lãi thu được nhờ bán xe trong khoảng thời gian trên là:

$$5 \cdot (0 \cdot 2 + 1 \cdot 13 + 2 \cdot 15 + 3 \cdot 12 + 4 \cdot 7 + 5 \cdot 3) = 610$$

Chi phí thuê nhân viên và cơ sở vật chất trong thời gian đó là:

$$2 \cdot (2 + 13 + 15 + 12 + 7 + 3) = 104$$

Vậy trong khoảng thời gian đó cửa hàng có lãi $610 - 104 = 506$ triệu đồng.

Bài 6. Điều tra về số tiền mua sách (đơn vị: nghìn đồng) trong một năm của 50 sinh viên, người ta thu được bảng số liệu thống kê sau

203	37	141	43	55	303	252	758	321	123
425	27	72	87	215	358	521	863	284	279
608	302	703	68	149	327	127	125	489	234
498	968	350	57	75	503	712	440	185	404
98	552	101	612	333	451	901	875	789	202

Từ bảng số liệu thống kê trên, người ta lập bảng phân bố tần số ghép lớp như sau

Lớp	Tần số
[0; 99]	10
[100; 199]	7
[200; 299]	7
[300; 399]	7
[400; 499]	6
[500; 599]	3
[600; 699]	2
[700; 799]	4
[800; 899]	2
[900; 999]	2
Cộng	$n = 50$

Xét tập 20% sinh viên dùng nhiều tiền để mua sách nhất. Người mua ít nhất trong nhóm này mua hết bao nhiêu tiền?

Lời giải. Xét tập 20% số sinh viên mua nhiều tiền nhất. Nhóm này có $50 \cdot 20\% = 10$ sinh viên. Có 10 sinh viên tiêu từ 600 nghìn đồng trở lên. Do bài toán hỏi người mua ít nhất nên ta xét trong nhóm [600; 699]. Nhóm này có hai người mua hết 608 nghìn đồng và 612 nghìn đồng. Do đó, người mua ít nhất là 608 nghìn đồng.

Bài 7. Một học sinh ghi lại bảng phân bố tần số, tần suất ghép lớp của một mẫu số liệu như sau

Lớp	[1; 9]	[10; 19]	[20; 29]	[30; 39]	[40; 49]	
Tần số						$n =$
Tần suất (%)	12,5	0,0	50,0	25,0	12,5	100

Tuy nhiên, em đó quên ghi kích thước mẫu n . Biết rằng n là số có 3 chữ số và chữ số tận cùng là 8. Tìm giá trị nhỏ nhất của n .

Lời giải. Hai lớp $[1; 9]$ và $[40; 49]$ có tần số là $n.12,5\% = \frac{n}{8}$.

Lớp $[20; 29]$ có tần số là $n.50\% = \frac{n}{2}$.

Lớp $[30; 39]$ có tần số là $n.25\% = \frac{n}{4}$.

Vì tần số là các số nguyên dương nên n phải chia hết cho 8; 4; 2. Mà n là số có 3 chữ số, chữ số tận cùng là 8 và nhỏ nhất nên $n = 128$.

Bài 8. Một cảnh sát giao thông ghi tốc độ (đơn vị: km/h) của 30 chiếc xe qua trạm như sau

53	47	59	66	36	69	83	77	42	57	51	60	78	63	46
63	42	55	63	48	75	60	58	80	44	59	60	75	49	63

Hãy lập bảng phân bố tần số - tần suất ghép lớp gồm 6 lớp với độ dài mỗi đoạn của lớp là 7.

Lời giải.

Lớp	Tần số	Tần suất (%)
$[36; 43]$	10	10
$[44; 51]$	7	20
$[52; 59]$	7	20
$[60; 67]$	7	26,7
$[68; 75]$	6	10
$[76; 83]$	3	13,3
Cộng	$n = 30$	100

BÀI TẬP TỔNG HỢP

Bài 1. Điều tra thời gian hoàn thành một sản phẩm của 20 công nhân, người ta thu được mẫu số liệu sau (thời gian tính bằng phút).

10	12	13	15	11	13	16	18	19	21
23	21	15	17	16	15	20	13	16	11

Kích thước mẫu là bao nhiêu?

Lời giải. Kích thước mẫu bằng 20

Bài 2. Điểm kiểm tra học kì môn Toán của các học sinh lớp 10A cho ở bảng dưới đây.

Điểm	3	4	5	6	7	8	9	10
Tần số	1	2	5	8	6	10	7	2

Hỏi lớp 10A có bao nhiêu học sinh?

Lời giải. Lớp 10A có 41 học sinh

Bài 3. Dưới đây là bảng phân bố tần số - tần suất của đại lượng X . Trong bảng còn hai số chưa biết x và y . Tìm x và y .

Dấu hiệu	9	10	12	15	16
Tần số	1	2	3	x	4
Tần suất (%)	5	10	15	50	y

Lời giải. Ta có $x = \frac{3.50}{15} = 10$; $y = \frac{15.4}{3} = 20$.

Bài 4.

Cho bảng phân bố tần số ghép lớp của một mẫu số liệu mà các số liệu được sắp xếp từ nhỏ đến lớn như bên. Hãy tính tần suất (%) của lớp chứa số liệu thứ 5 của mẫu số liệu.

Lớp	Tần số
[0; 3]	3
[4; 7]	1
[7; 10]	2
[8; 11]	4

Lời giải. Số liệu thứ 5 thuộc lớp [7; 10]. Do đó, tần suất cần tìm là $\frac{2}{10} = 20\%$

Bài 5. Cho bảng số liệu về khối lượng của 30 củ khoai tây thu hoạch từ một thửa ruộng như dưới đây.

Lớp khối lượng (gam)	Tần số
[70; 80)	3
[80; 90)	6
[90; 100)	12
[100; 110)	6
[110; 120]	3
Cộng	30

Tần suất của lớp [100; 110) là bao nhiêu?

Lời giải. Tần suất ghép lớp [100; 110) là $\frac{6}{30} \cdot 100\% = 20\%$

Bài 6. Kết quả khảo sát ở 43 tỉnh (đơn vị: %) ghi lại số phần trăm những trẻ em mới sinh có trọng lượng dưới 2500 gam được cho ở bảng phân bố tần số ghép lớp bên dưới.

Lớp	[4, 5; 5, 4]	[5, 5; 6, 4]	[6, 5; 7, 4]	[7, 5; 8, 4]	[8, 5; 9, 4]	Cộng
Tần số	9	6	17	8	3	$n = 43$

Tính tỷ lệ phần trăm số tỉnh có số phần trăm những trẻ em mới sinh có trọng lượng dưới 2500 gam không vượt quá 7,4% (làm tròn một chữ số thập phân).

Lời giải. Tỷ lệ phần trăm số tỉnh có số phần trăm những trẻ em mới sinh có trọng lượng dưới 2500 gam không vượt quá 7,4% là $\frac{9+6+17}{35} \approx 74,41\%$

Bài 7.

Người ta thống kê số phần trăm trẻ em mới sinh có trọng lượng dưới 2,5 kg của 43 tỉnh thành ở bảng phân bố tần suất ghép lớp sau. Có bao nhiêu tỉnh có tỉ lệ phần trăm trẻ có trọng lượng dưới 2,5 kg dưới 7,5% ?

Lớp	Tần suất (%)
[4, 5; 5, 5)	9
[5, 5; 6, 5)	6
[6, 5; 7, 5)	17
[7, 5; 8, 5)	8
[8, 5; 9, 5)	3

Lời giải. Chưa đủ giả thiết để kết luận

Bài 8. Một học sinh ghi lại bảng phân bố tần số, tần suất ghép lớp của một mẫu số liệu như sau

Lớp	[1; 9]	[10; 19]	[20; 29]	[30; 39]	[40; 49]	
Tần số						$n =$
Tần suất (%)	12,5	0,0	50,0	25,0	12,5	100

Tuy nhiên, em đó quên ghi kích thước mẫu n . Tìm giá trị nhỏ nhất của n .

Lời giải. Lớp [1; 9] có tần số là $n \cdot 12,5\% = \frac{n}{8}$.

Lớp [20; 29] có tần số là $n \cdot 50\% = \frac{n}{2}$.

Lớp [30; 39] có tần số là $n \cdot 25\% = \frac{n}{4}$.

Lớp $[40; 49]$ có tần số là $n.12,5\% = \frac{n}{8}$.

Vì tần số là các số nguyên dương nên n phải chia hết cho 8;4;2. Do đó, số n nhỏ nhất thỏa có giá trị bằng 8.

§2. BIỂU ĐỒ

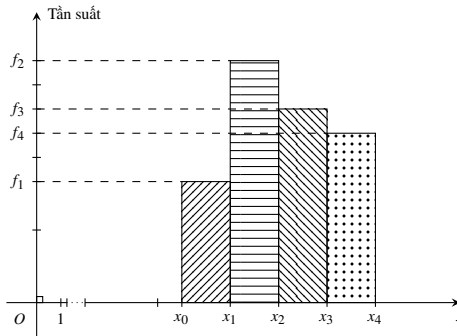
I. Tóm tắt lí thuyết

1. Biểu đồ tần suất hình cột

Cho bảng phân bố tần suất ghép lớp

Lớp x	$[x_0; x_1)$	$[x_1; x_2)$	$[x_2; x_3)$	$[x_3; x_4]$	Cộng
Tần suất (%)	f_1	f_2	f_3	f_4	100 %

Ta có thể mô tả bảng phân bố tần suất ghép lớp qua biểu đồ tần suất hình cột như hình vẽ sau



⚠ Tương tự, ta cũng có thể vẽ biểu đồ hình cột mô tả cho bảng phân bố tần suất, tần số và tần số ghép lớp.

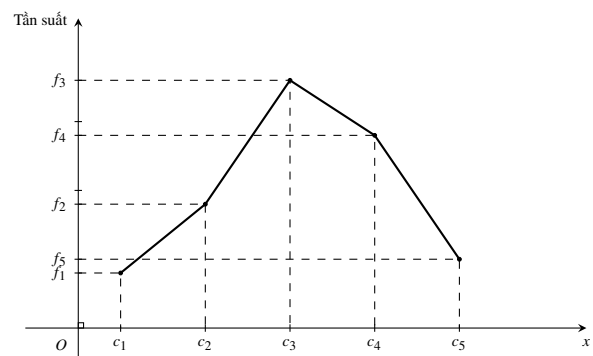
2. Đường gấp khúc tần suất

Cho bảng phân bố tần suất ghép lớp

Lớp x	$[x_0; x_1)$	$[x_1; x_2)$	$[x_2; x_3)$	$[x_3; x_4)$	$[x_4; x_5]$	Cộng
Tần suất (%)	f_1	f_2	f_3	f_4	f_5	100 %

Ta cũng có thể mô tả bảng tần suất ghép lớp bằng một đường gấp khúc qua các bước sau:

- Trên mặt phẳng tọa độ ta xác định các điểm $(c_i; f_i)$, $i = 1, 2, 3, 4, 5$ trong đó c_i là trung bình cộng của hai đầu nút của lớp i (ta gọi c_i là giá trị đại diện của lớp i).
- Nối các điểm $(c_i; f_i)$ với $(c_{i+1}; f_{i+1})$, $i = 1, 2, 3, 4$, ta thu được đường gấp khúc, gọi là đường gấp khúc tần suất.



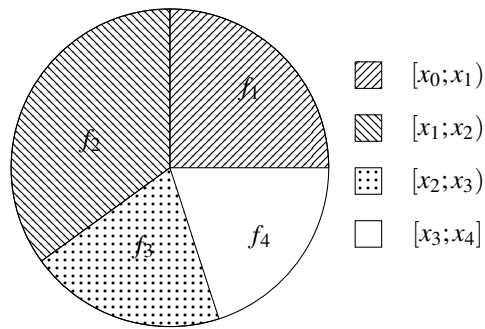
⚠ Ta cũng có thể mô tả bảng tần suất, tần số và tần số ghép lớp bằng cách vẽ biểu đồ hình cột hoặc đường gấp khúc. Ở đó ta chỉ cần thay cột tần suất bằng cột tần số.

3. Biểu đồ hình quạt

Cho bảng phân bố tần suất ghép lớp

Lớp x	$[x_0; x_1)$	$[x_1; x_2)$	$[x_2; x_3)$	$[x_3; x_4]$	Cộng
Tần suất (%)	f_1	f_2	f_3	f_4	100 %

Ta có thể mô tả bảng phân bố tần suất ghép lớp thông qua biểu đồ hình quạt như sau



! Chúng ta cũng có thể dùng biểu đồ hình quạt để mô tả bảng phân bố tần số, tần số ghép lớp.

II. Các dạng toán

Dạng 1. Vẽ biểu đồ tần số và tần suất hình cột

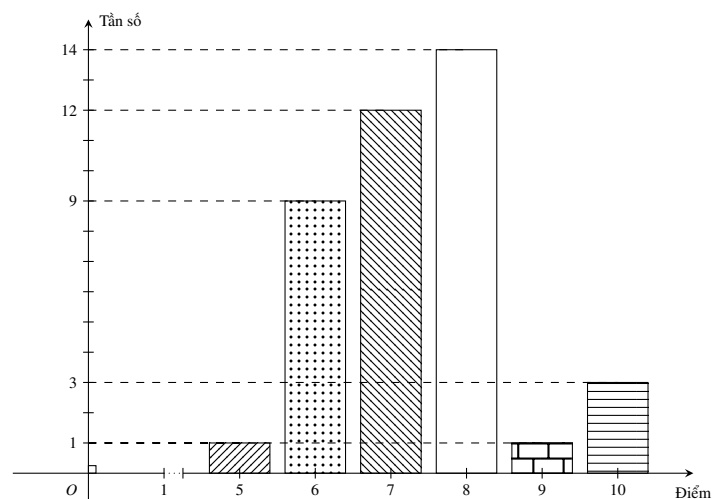
Từ các số liệu trong bảng phân bố tần số và tần suất, ta tọa độ hóa để vẽ biểu đồ trên hệ trục tọa độ.

Ví dụ 1. Kết quả điểm thi môn Toán của lớp 10A9 trường THPT Đông Thụy Anh được cho theo bảng dưới đây.

Điểm	5	6	7	8	9	10	Cộng
Tần số	1	9	12	14	1	3	40

Hãy vẽ biểu đồ tần số và tần suất dạng cột để mô tả cho bảng số liệu trên.

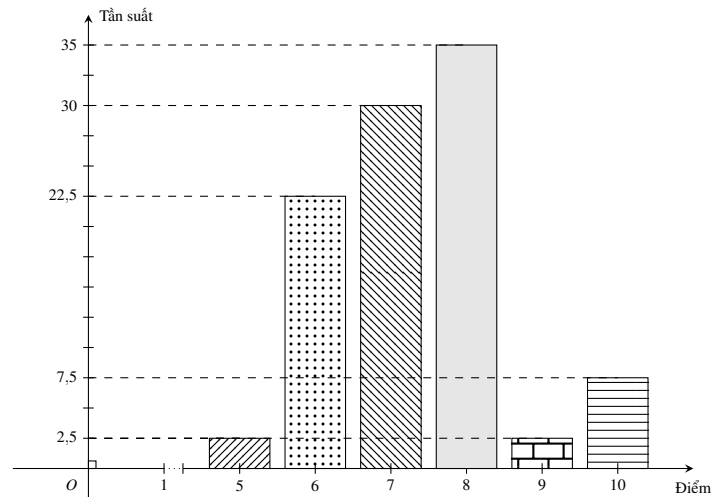
Lời giải. Biểu đồ tần số hình cột mô tả kết quả điểm thi của lớp 10A9



Bảng phân bố tần suất

Điểm	5	6	7	8	9	10	Cộng
Tần suất	2,5	22,5	30	35	2,5	7,5	100 %

Biểu đồ tần số hình cột mô tả kết quả điểm thi của lớp 10A9



Ví dụ 2. Số điện tiêu thụ của 30 hộ ở một khu dân cư trong một tháng được thống kê theo bảng sau:

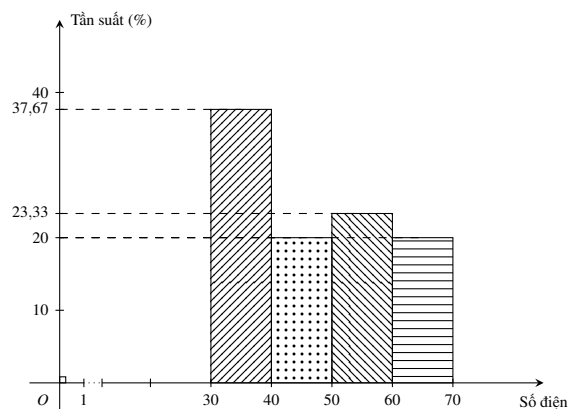
50	47	30	65	63	70	38	34	48	53
33	39	32	40	50	55	50	61	37	37
43	35	65	60	31	33	41	45	55	59

Hãy lập bảng phân bố tần số, tần suất ghép lớp và lập biểu đồ tần suất hình cột mô tả cho bảng phân bố tần suất ghép lớp đó.

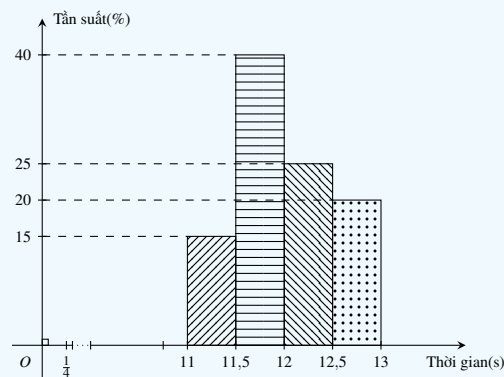
Lời giải. Bảng phân bố tần số và tần suất ghép lớp:

Lớp số điện (kW)	[30; 40)	[40; 50)	[50; 60)	[60; 70]	Cộng
Tần số	11	6	7	6	30
Tần suất (%)	37,67	20	23,33	20	100%

Biểu đồ tần suất ghép lớp hình cột mô tả cho bảng số liệu đã cho



Ví dụ 3. Cho biểu đồ tần suất ghép lớp hình cột mô tả thành tích bạn Bình tập chạy trong 20 lần như hình dưới đây. Hãy lập bảng phân bố tần suất ghép lớp mà biểu đồ đã mô tả.



Lời giải. Bảng phân bố tần suất thể hiện thành tích tập chạy của bạn Bình trong 20 lần:

Lớp thời gian chạy (s)	[11; 11,5)	[11,5; 12)	[12; 12,5)	[12,5; 13]	Cộng
Tần suất (%)	15	40	25	20	100%

BÀI TẬP TỰ LUYỆN

Bài 1. Một công ty sản xuất bóng đèn kiểm tra định kì bằng cách thắp thử nghiệm 30 bóng đèn để kiểm tra tuổi thọ (tính theo giờ). Kết quả của cuộc thử nghiệm được thống kê theo bảng sau:

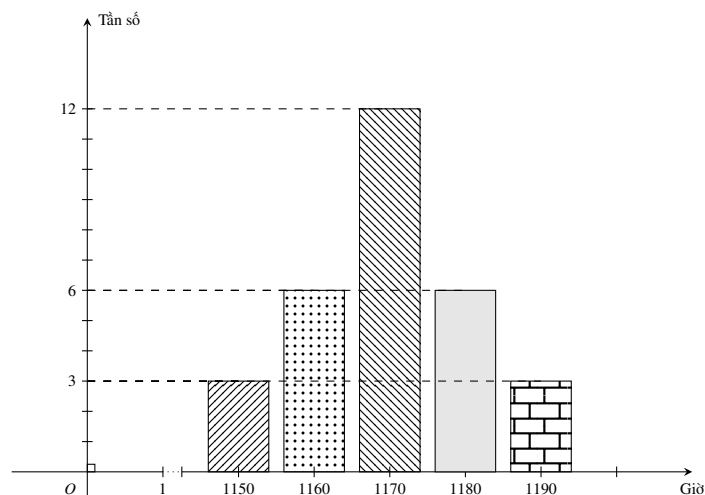
1180	1150	1190	1170	1180	1170	1160	1170	1160	1150
1190	1180	1170	1170	1170	1190	1170	1170	1170	1180
1170	1160	1160	1160	1170	1160	1180	1180	1150	1170

Hãy lập bảng phân bố tần số và vẽ biểu đồ tần số hình cột.

Lời giải. Bảng phân bố tần số

Giờ	1150	1160	1170	1180	1190	Cộng
Tần số	3	6	12	6	3	30

Mô tả bảng phân bố tần số bằng biểu đồ tần số hình cột



Bài 2. Số cuộn phim mà 40 nhà nhiếp ảnh nghiệp dư sử dụng trong một tháng được thống kê bằng bảng số liệu sau:

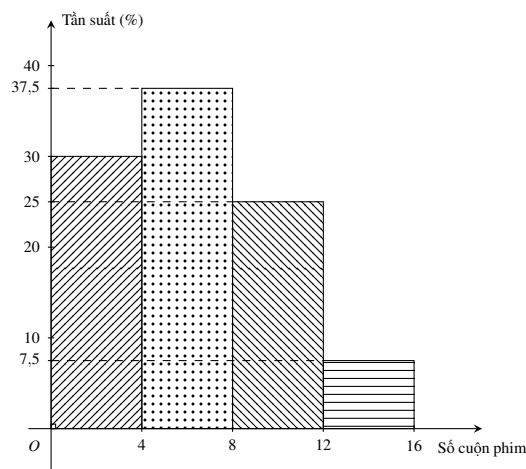
5	3	3	1	4	3	4	3	6	8
4	2	4	6	8	9	6	2	10	11
15	1	2	5	13	7	7	2	5	4
3	16	10	4	7	2	10	11	8	9

Hãy lập bảng tần số, tần suất ghép lớp và mô tả bảng phân bố tần suất ghép lớp bằng biểu đồ tần suất hình cột.

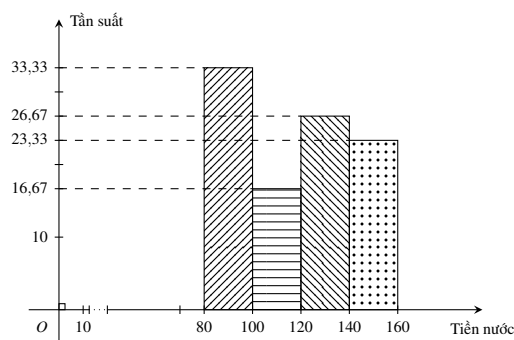
Lời giải. Bảng phân bố tần số và tần suất ghép lớp của bảng thống kê số cuộn phim của 40 nhà nhiếp ảnh nghiệp dư sử dụng trong một tháng là

Lớp số cuộn phim	[0; 4)	[4; 8)	[8; 12)	[12; 16]	Cộng
Tần số	12	15	10	3	40
Tần suất (%)	30	37,5	25	7,5	100%

Biểu đồ mô tả cho bảng tần suất ghép lớp



Bài 3. Cho biểu đồ tần suất ghép lớp hình cột mô tả thống kê tiền nước (nghìn đồng) phải trả hàng tháng của gia đình anh Huy trong năm 2017 như hình vẽ:



Hãy lập bảng phân bố tần suất ghép lớp về tiền điện hàng tháng phải trả trong năm 2017 của gia đình anh Huy.

Lời giải. Bảng phân bố tần suất ghép lớp tiền nước hàng tháng của gia đình anh Huy trong năm 2017 như sau:

Tiền nước(nghìn đồng)	[80; 100)	[100; 120)	[120; 140)	140; 160	Cộng
Tần suất(%)	33,33	16,67	26,67	23,33	100%

Dạng 2. Biểu đồ đường gấp khúc

Ví dụ 4. Cho bảng tần suất ghép lớp:

Điểm thi học kì I môn Toán của 40 học sinh lớp 10D3 của trường THPT A

Lớp điểm thi	[0; 2)	[2; 4)	[4; 6)	[6; 8)	[8; 10]	Cộng
Tần suất (%)	7,5	12,5	40	30	10	100%

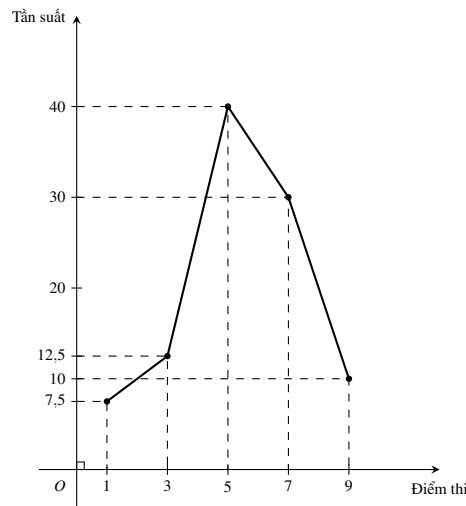
- Hãy vẽ biểu đồ đường gấp khúc tần suất về điểm thi học kì I môn Toán của 40 học sinh lớp 10D3 của trường THPT A.
- Có bao nhiêu học sinh thuộc lớp điểm chiếm tỉ lệ cao nhất.
- Biết điểm giỏi là từ 8 đến 10. Hỏi lớp 10D3 có bao nhiêu học sinh đạt điểm giỏi.

Lời giải.

- a) Bảng giá trị đại diện (GTĐD) của các lớp:

Lớp điểm thi	[0; 2)	[2; 4)	[4; 6)	[6; 8)	[8; 10]
GTĐD	1	3	5	7	9

Biểu đồ đường gấp khúc tần suất về điểm thi học kì I môn Toán của 40 học sinh lớp 10D3 của trường THPT A:



- Lớp điểm $[4; 6)$ chiếm tỉ lệ cao nhất bằng 40 %. Suy ra số học sinh thuộc lớp này bằng: $\frac{40 \times 40}{100} = 16$.
- Lớp điểm $[8; 10]$ chiếm tỉ lệ 10 % nên số học sinh đạt điểm giỏi là: $\frac{10 \times 40}{100} = 4$.

Ví dụ 5. Cho các bảng tần số ghép lớp:

Chiều cao (cm) của 40 học sinh lớp 10A1 của trường THPT B

Lớp chiều cao	[1,5; 1,55)	[1,55; 1,6)	[1,6; 1,65)	[1,65; 1,7)	[1,7; 1,75]	Cộng
Tần số	0	2	15	18	5	40

Chiều cao (cm) của 40 học sinh lớp 10D1 của trường THPT B

Lớp chiều cao	[1,5; 1,55)	[1,55; 1,6)	[1,6; 1,65)	[1,65; 1,7)	[1,7; 1,75]	Cộng
Tần số	1	12	23	4	0	40

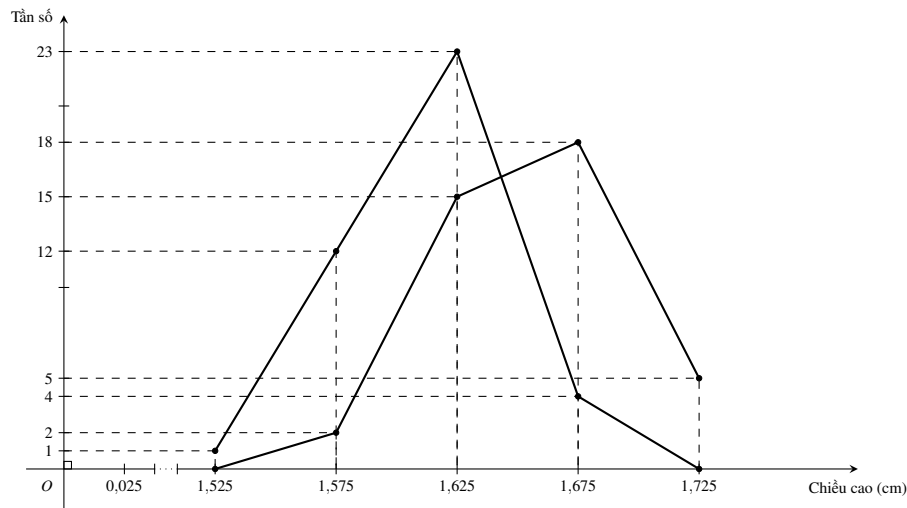
- Hãy vẽ đường gấp khúc tần số về chiều cao của 40 học sinh lớp 10A1 và đường gấp khúc tần số về chiều cao của 40 học sinh lớp 10D1 của trường THPT B trên cùng một biểu đồ.
- Nhận xét về chiều cao trung bình của học sinh lớp 10A1 so với lớp 10D1.

Lời giải.

a) Bảng giá trị đại diện (GTĐĐ) của các lớp:

Lớp chiều cao	[1,5; 1,55)	[1,55; 1,6)	[1,6; 1,65)	[1,65; 1,7)	[1,7; 1,75]
GTĐĐ	1,525	1,575	1,625	1,675	1,725

Biểu đồ đường gấp khúc tần số về chiều cao của 40 học sinh lớp 10A1 và đường gấp khúc tần số về chiều cao của 40 học sinh lớp 10D1 của trường THPT B:



b) Nhận xét: Sĩ số hai lớp bằng nhau. Đường gấp khúc biểu diễn chiều cao từ 1,65 cm trở lên của lớp 10A1 nằm trên lớp 10D1, đường gấp khúc biểu diễn chiều cao dưới 1,65 cm của lớp 10A1 nằm dưới lớp 10D1. Vậy chiều cao trung bình lớp 10A1 lớn hơn lớp 10D1.

BÀI TẬP TỰ LUYỆN

Bài 4. Cho bảng tần suất ghép lớp:

Cân nặng (kg) của 50 học sinh lớp 10A3 của trường THPT C

Lớp cân nặng	[40; 42)	[42; 44)	[44; 46)	[46; 48)	[48; 50)	[50; 52]	Cộng
Tần suất (%)	2	10	24	38	20	6	100 %

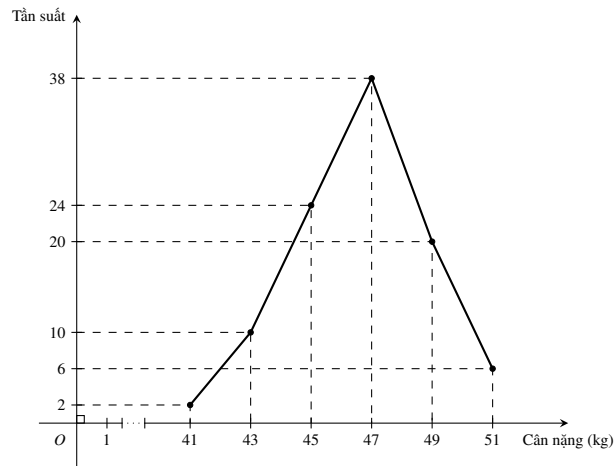
- a) Hãy vẽ biểu đồ đường gấp khúc tần suất về cân nặng của 50 học sinh lớp 10A3 của trường THPT C.
- b) Có bao nhiêu học sinh có cân nặng từ 48 kg đến 52 kg.

Lời giải.

1. Bảng giá trị đại diện (GTĐĐ) của các lớp:

Lớp cân nặng	[40; 42)	[42; 44)	[44; 46)	[46; 48)	[48; 50)	[50; 52]
GTĐĐ	41	43	45	47	49	51

Biểu đồ đường gấp khúc tần suất về cân nặng của 50 học sinh lớp 10A3 của trường THPT C:



2. Học sinh có cân nặng từ 48 kg đến 52 kg chiếm tỉ lệ $20\% + 6\% = 26\%$. Suy ra số học sinh có cân nặng từ 48 kg đến 52 kg là $\frac{26 \times 50}{100} = 13$.

Bài 5. Cho các bảng tần số ghép lớp:

Điểm phẩy học kì 1 môn Toán của 40 học sinh lớp 10A1 của trường THPT A

Lớp điểm phẩy	[5; 6)	[6; 7)	[7; 8)	[8; 9)	[9; 10]	Cộng
Tần số	0	2	25	10	3	40

Điểm phẩy học kì 1 môn Văn của 40 học sinh lớp 10A1 của trường THPT A

Lớp điểm phẩy	[5; 6)	[6; 7)	[7; 8)	[8; 9)	[9; 10]	Cộng
Tần số	5	23	11	1	0	40

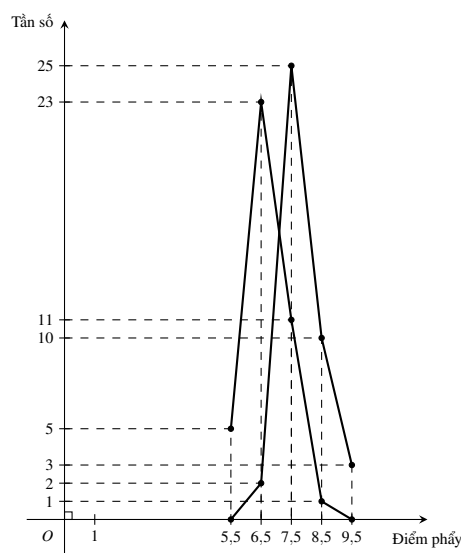
- a) Hãy vẽ đường gấp khúc tần số về điểm phẩy học kì 1 môn Toán và đường gấp khúc tần số về điểm phẩy học kì 1 môn Văn của 40 học sinh lớp 10A1 trên cùng một biểu đồ.
- b) Nhận xét về điểm phẩy trung bình môn Toán so với môn Văn của học sinh lớp 10A1.

Lời giải.

1. Bảng giá trị đại diện (GTĐĐ) của các lớp:

Lớp điểm phẩy	[5; 6)	[6; 7)	[7; 8)	[8; 9)	[9; 10]
GTĐĐ	5,5	6,5	7,5	8,5	9,5

Biểu đồ đường gấp khúc tần số về điểm phẩy học kì 1 môn Toán và đường gấp khúc tần số về điểm phẩy học kì 1 môn Văn của 40 học sinh lớp 10A1:



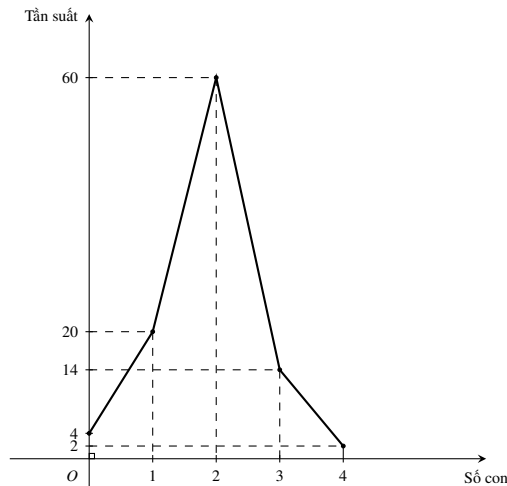
2. Nhận xét: Đường gấp khúc biểu diễn điểm phẩy từ 7,0 trở lên của môn Toán nằm trên môn Văn, đường gấp khúc biểu diễn điểm phẩy dưới 7,0 của môn Toán nằm dưới môn Văn. Vậy điểm phẩy trung bình môn Toán cao hơn môn Văn.

Bài 6. Cho bảng tần suất:

Số con của 50 hộ gia đình ở địa phương A						
Số con	0	1	2	3	4	Cộng
Tần suất (%)	4	20	60	14	2	100 %

Hãy vẽ biểu đồ đường gấp khúc tần suất về số con của 50 hộ gia đình ở địa phương A và nhận xét xem có bao nhiêu hộ gia đình chưa thực hiện tốt kế hoạch hóa gia đình (có nhiều hơn 2 con).

Lời giải. Biểu đồ đường gấp khúc tần suất về số con của 50 hộ gia đình ở địa phương A:



Các gia đình có nhiều hơn 2 con chiếm $14\% + 2\% = 16\%$. Suy ra số gia đình có nhiều hơn 2 con là $\frac{16 \times 50}{100} = 8$.

Bài 7. Cho bảng tần suất ghép lớp:

Tốc độ (km/h) của 40 chiếc xe máy qua trạm kiểm soát giao thông B

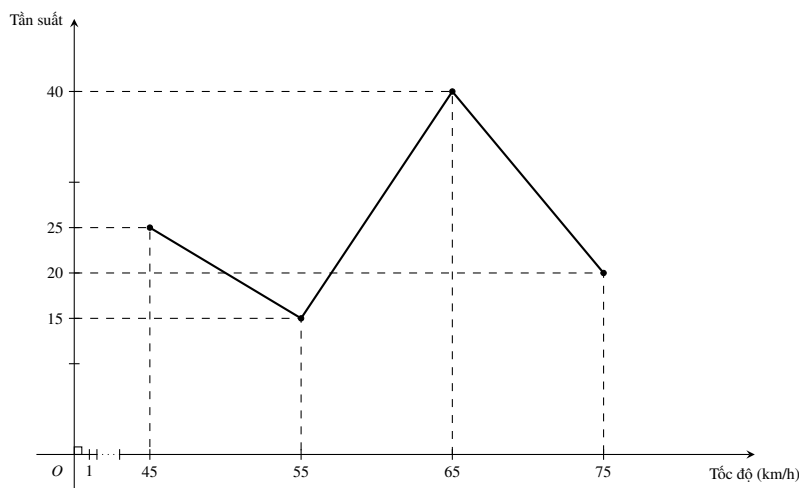
Lớp tốc độ	[40; 50)	[50; 60)	[60; 70)	[70; 80]	Cộng
Tần suất (%)	25	15	40	20	100 %

Hãy vẽ biểu đồ đường gấp khúc tần suất về tốc độ (km/h) của 40 chiếc xe máy qua trạm kiểm soát giao thông B và nhận xét xem có bao nhiêu chiếc xe đi với tốc độ không dưới 70 km/h.

Lời giải. Bảng giá trị đại diện (GTĐĐ) của các lớp:

Lớp tốc độ	[40; 50)	[50; 60)	[60; 70)	[70; 80]
GTĐĐ	45	55	65	75

Biểu đồ đường gấp khúc tần suất về tốc độ (km/h) của 40 chiếc xe máy qua trạm kiểm soát giao thông B:



Các xe đi với vận tốc không dưới 70 km/h chiếm 20 %. Suy ra số xe đi với vận tốc không dưới 70 km/h là $\frac{20 \times 40}{100} = 8$.

Bài 8. Cho bảng tần suất:

Tiền lương (nghìn đồng) hàng tháng của 50 công nhân ở xưởng may C

Tiền lương	700	800	900	1000	1100	1200	Cộng
Tần suất (%)	16	18	24	20	12	10	100 %

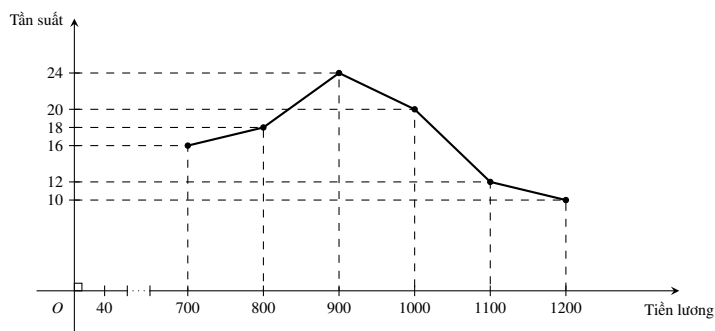
Hãy vẽ biểu đồ đường gấp khúc tần suất và đường gấp khúc tần số về tiền lương (nghìn đồng) hàng tháng của 50 công nhân ở xưởng may C.

Lời giải. Bảng phân bố tần số:

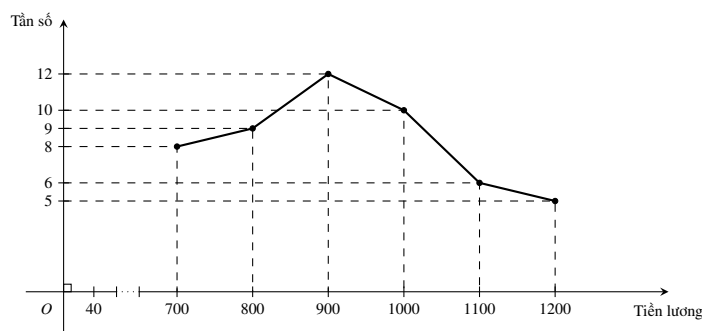
Tiền lương (nghìn đồng) hàng tháng của 50 công nhân ở xưởng may C

Tiền lương	700	800	900	1000	1100	1200	Cộng
Tần số	8	9	12	10	6	5	50

Biểu đồ đường gấp khúc tần suất về tiền lương (nghìn đồng) hàng tháng của 50 công nhân ở xưởng may C:



Biểu đồ đường gấp khúc tần suất về tiền lương (nghìn đồng) hàng tháng của 50 công nhân ở xưởng may C:



Dạng 3. Biểu đồ hình quạt

Ví dụ 6. Chiều cao (cm) của 36 học sinh nam ở một lớp 12A1:

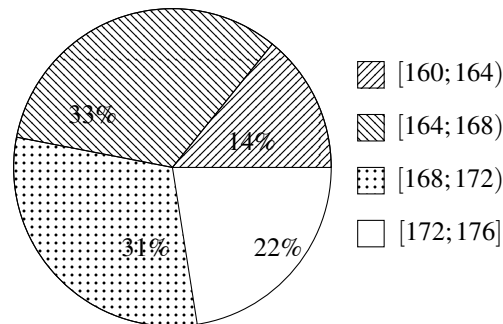
Lớp chiều cao (cm)	[160; 164)	[164; 168)	[168; 172)	[172; 176]	Cộng
Tần số	5	12	11	8	36

Vẽ biểu đồ tần suất hình quạt cho bảng thống kê trên.

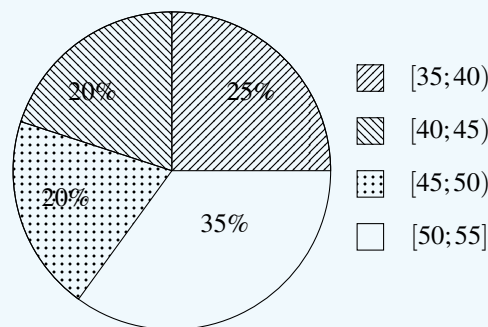
Lời giải. Ta có bảng phân bố tần suất sau

Lớp chiều cao (cm)	[160; 164)	[164; 168)	[168; 172)	[172; 176]	Cộng
Tần số	5	12	11	8	36
Tần suất (%)	14	33	31	22	100

Từ đó ta có biểu đồ tần suất hình quạt như sau:



Ví dụ 7. Tiến hành một cuộc thăm dò về số cân nặng của một nhóm xã. Kết quả thu được biểu diễn qua biểu đồ tần suất hình quạt như sau:



Hỏi số người trong xã có cân nặng từ 40 kg đến dưới 50 kg là bao nhiêu người, biết rằng xã có 1000 người.

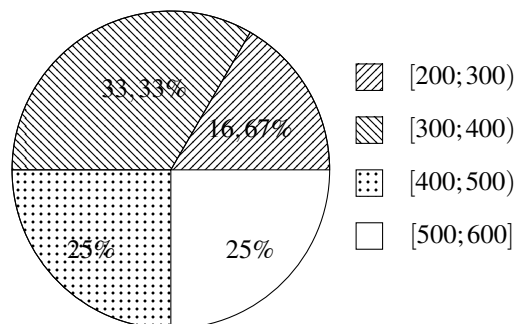
Lời giải. Dựa vào biểu đồ quạt ta có:

Tỉ lệ người có cân nặng từ 40 kg đến dưới 50 kg là $20 + 20 = 40 \%$.

Số người có cân nặng từ 40 kg đến dưới 50 kg là $\frac{40 \times 1000}{100} = 400$ người.

BÀI TẬP TỰ LUYỆN

Bài 9. Số lượng khách đến tham quan tại Đà Nẵng trong 12 tháng được cho bởi biểu đồ như sau:

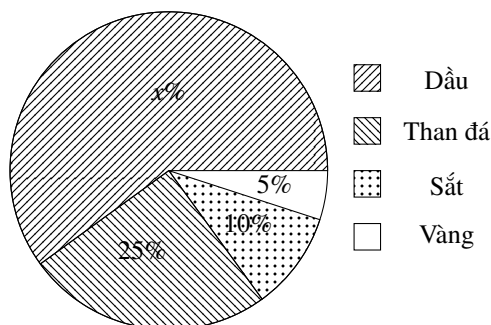


Tính số tháng mà số người tham quan không dưới 400 người.

Lời giải. Dựa vào biểu đồ, ta có số người tham quan không dưới 400 người chiếm tỉ lệ phần trăm là: $25 + 25 = 50 \%$.

Vậy số tháng mà số người thăm quan trên 400 người là $\frac{50 \times 12}{100} = 6$ tháng.

Bài 10. Biểu đồ hình quạt sau mô tả tỉ lệ về giá trị đạt được của khoáng sản xuất khẩu nước ngoài của nước ta:



Biết rằng giá trị xuất khẩu của nước ta về dầu là 450 triệu USA. Hỏi giá trị xuất khẩu vàng là bao nhiêu triệu USA?

Lời giải. Tỉ lệ phần trăm của dầu là $100\% - 25\% - 10\% - 5\% = 60\%$.

Suy ra giá trị xuất khẩu của vàng là $\frac{450}{60} \cdot 5 = 37,5$ triệu USA.

Bài 11. Cho bảng phân bố tần số điểm thi môn Anh Văn của một trung tâm ở Hà Nội:

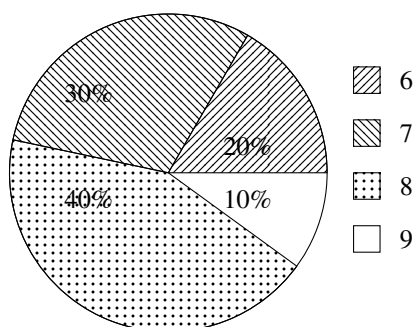
Điểm thi	6	7	8	9	Cộng
Tần số	40	60	80	20	200

Vẽ biểu đồ hình quạt mô tả bảng dữ liệu thống kê trên.

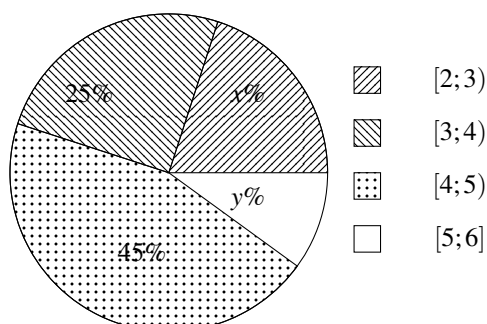
Lời giải. Ta có bảng phân bố tần suất sau

Điểm thi	6	7	8	9	Cộng
Tần số	40	60	80	20	200
Tần suất (%)	20	30	40	10	100

Khi đó ta có biểu đồ tần suất hình quạt:



Bài 12. Tuổi thọ (tính theo tháng) của 100 bóng đèn thấp thử được thể hiện qua biểu đồ tần suất hình quạt:



Biết rằng x, y thỏa mãn $x^2 + y^2 = 500$. Tìm số bóng đèn có tuổi thọ trong khoảng $[2; 5)$ biết rằng $x > 15$.

Lời giải. Ta có hệ

$$\begin{cases} x + y = 30 \\ x^2 + y^2 = 500 \end{cases} \Leftrightarrow \begin{cases} y = 30 - x \\ x^2 + (30 - x)^2 = 500 \end{cases} \Leftrightarrow \begin{cases} y = 30 - x \\ 2x^2 - 60x + 400 = 0 \end{cases} \Leftrightarrow \begin{cases} y = 30 - x \\ \begin{cases} x = 20 \\ x = 10 \end{cases} \end{cases} \Leftrightarrow \begin{cases} x = 20 \\ y = 10 \end{cases}.$$

Tỉ lệ của số bóng đèn có tuổi thọ trong khoảng $[2; 5)$ là $20\% + 25\% + 45\% = 90\%$. Khi đó, số bóng đèn có tuổi thọ trong khoảng $[2; 5)$ là $\frac{90 \times 100}{100} = 90$ bóng.

§3. SỐ TRUNG BÌNH CỘNG. SỐ TRUNG VỊ. MỘT

I. Tóm tắt lý thuyết

1. Số trung bình cộng

Định nghĩa 1 (Số trung bình cộng). Số trung bình cộng (số trung bình) của một dãy gồm n số liệu x_1, x_2, \dots, x_n kí hiệu là \bar{x} và được tính theo công thức: $\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$

- Trường hợp bảng phân bố tần số và tần suất

Giá trị	x_1	x_2	\dots	x_m	Cộng
Tần số	n_1	n_2	\dots	n_m	N
Tần suất	f_1	f_2	\dots	f_m	100%

Số trung bình cộng được tính theo công thức:

$\bar{x} = \frac{n_1x_1 + n_2x_2 + \dots + n_kx_k}{n} = f_1x_1 + f_2x_2 + \dots + f_kx_k$ trong đó n_i, f_i lần lượt là tần số, tần suất của giá trị x_i ($i = 1, 2, \dots, k$) và n là số các số liệu thống kê ($n = n_1 + n_2 + \dots + n_k$).

- Trường hợp bảng phân bố tần số và tần suất ghép lớp


Lớp	Giá trị đại diện	Tần số
$[a_1; a_2)$	x_1	n_1
$[a_2; a_3)$	x_2	n_2
\dots	\dots	\dots
$[a_{m-1}; a_m)$	x_m	n_m
		$N = \sum_{i=1}^m n_i$

Số trung bình cộng được tính theo công thức:

$\bar{x} = \frac{n_1c_1 + n_2c_2 + \dots + n_kc_k}{n} = f_1c_1 + f_2c_2 + \dots + f_kc_k$ trong đó n_i, c_i, f_i lần lượt là giá trị đại diện, tần số, tần suất của lớp thứ i ($i = 1, 2, \dots, k$) và n là số các số liệu thống kê ($n = n_1 + n_2 + \dots + n_k$).

2. Số trung vị


Định nghĩa 2 (Số trung vị). Sắp thứ tự các số liệu thống kê thành dãy không giảm (hoặc không tăng). Số trung vị (của các số liệu thống kê đã cho) kí hiệu M_e là số đứng giữa dãy nếu số phần tử là lẻ và là trung bình cộng của hai số đứng giữa dãy nếu số phần tử là chẵn.

 Số trung vị được xác định như sau:

- $M_e = x_{\frac{n+1}{2}}$ nếu n là số lẻ.
- $M_e = \frac{1}{2} \left(x_{\frac{n}{2}} + x_{\frac{n+1}{2}} \right)$ n là số chẵn.

3. Một

Định nghĩa 3 (Một). Một của một bảng phân bố tần số là giá trị có tần số lớn nhất và được kí hiệu là M_O .

 Chú ý:

- Số trung bình của mẫu số liệu được dùng làm đại diện cho các số liệu của mẫu.

- Nếu các số liệu trong mẫu có sự chênh lệch quá lớn thì dùng số trung vị làm đại diện cho các số liệu của mẫu.
- Nếu quan tâm đến giá trị có tần số lớn nhất thì dùng mode làm đại diện. Một bảng phân bố tần số có thể có hai hay nhiều mode.

II. Các dạng toán

Dạng 1. Số trung bình

Áp dụng công thức số trung bình cho bảng số tần số, tần suất và tần số, tần suất ghép lớp.

Ví dụ 1. Khối lượng 30 chi tiết máy được cho bởi bảng sau

Khối lượng(gam)	250	300	350	400	450	500	Cộng
Tần số	4	4	5	6	4	7	30

Tính số trung bình \bar{x} (làm tròn đến chữ số thứ hai sau dấu phẩy) của bảng nói trên.

Lời giải. Áp dụng công thức tính số trung bình cho bảng tần số ta có

$$\bar{x} = \frac{250.4 + 300.4 + 350.5 + 400.6 + 450.4 + 500.7}{30} \approx 388,33 \text{ (gam)}.$$

Ví dụ 2. Chiều cao của 20 cây giống được cho bởi bảng sau:

Lớp (cm)	Tần số
[40; 44]	2
[45; 49]	5
[50; 54]	3
[55; 59]	4
[60; 64]	3
[65; 69]	3
	$N = 20$

Tính số trung bình \bar{x} (làm tròn đến chữ số thứ hai sau dấu phẩy) của bảng nói trên.

Lời giải. Bảng tần số ghép lớp của bảng nói trên là

Lớp (cm)	Giá trị đại diện	Tần số
[40; 44]	42	2
[45; 49]	47	5
[50; 54]	52	3
[55; 59]	57	4
[60; 64]	62	3
[65; 69]	67	3
		$N = 20$

Áp dụng công thức tính số trung bình cho bảng tần số ghép lớp ta có

$$\bar{x} \approx \frac{42.2 + 47.5 + 52.3 + 57.4 + 62.3 + 67.3}{20} = 54.5 \text{ (cm)}.$$

Dạng 2. Số trung vị

Áp dụng định nghĩa của số trung vị. Lưu ý có hai trường hợp khác nhau là mẫu số liệu có kích thước lẻ và mẫu số liệu có kích thước chẵn.

Ví dụ 3. Điều tra số học sinh của 30 lớp học, ta được bảng số liệu như sau:

35	39	39	40	40	41	41	41	41	44	44	45	45	45	46
48	48	48	48	49	49	49	49	49	49	50	50	50	50	51

Tính số trung vị của bảng nói trên.

Lời giải. Ta có $N = 30$ là số chẵn. Số liệu thứ 15 và 16 lần lượt là 46, 48. Vậy số trung vị là $M_e = \frac{46 + 48}{2} = 47$ (Học sinh).

Ví dụ 4. Điểm học kì một của một học sinh được cho bởi bảng số liệu sau (Đơn vị: điểm)

5	6	6	7	7	8	8	8,5	9
---	---	---	---	---	---	---	-----	---

Tính số trung vị của bảng nói trên.

Lời giải. Ta có $N = 9$ là số lẻ. Số liệu thứ $\frac{N+1}{2} = 5$ là số trung vị. Do đó số trung vị là $M_e = 7$ (Điểm).

BÀI TẬP TỰ LUYỆN

Bài 1. Bảng số liệu sau đây thống kê thời gian nảy mầm một loại hạt mới trong các điều kiện khác nhau

Thời gian(phút)	420	440	450	480	500	540
Tần số	8	17	18	16	11	10

Tính giá trị trung bình \bar{x} (làm tròn đến hai chữ số sau dấu phẩy) về thời gian nảy mầm loại hạt mới nói trên.

Lời giải. Áp dụng công thức tính số trung bình cho bảng tần số ta có

$$\bar{x} = \frac{420 \cdot 8 + 440 \cdot 17 + 450 \cdot 18 + 480 \cdot 16 + 500 \cdot 11 + 540 \cdot 10}{80} = 469 \text{ (Phút)}.$$

Bài 2. Điều tra số học sinh giỏi khối 10 của 15 trường cấp ba trên địa bàn tỉnh A, ta được bảng số liệu như sau:

22	29	29	29	30	31	32	32	33	34	34	35	35	35	36
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

Tính số trung vị của bảng nói trên.

Lời giải. Ta có $N = 15$ là số lẻ. Số liệu thứ $\frac{15+1}{2} = 8$. Vậy số trung vị là $M_e = 32$ (Học sinh).

Bài 3. Tốc độ phát triển của một loại Vi-rút trong 10 ngày với các điều kiện khác nhau (đơn vị nghìn con) được thống kê như sau

20	100	30	980	440	20	20	150	60	270
----	-----	----	-----	-----	----	----	-----	----	-----

Trong trường hợp này ta chọn số nào dưới đây làm giá trị đại diện là tốt nhất? Tính giá trị đại diện đó.

Lời giải. Ta chọn số trung vị làm đại diện là tốt nhất vì có sự chênh lệch lớn giữa các số liệu trong mẫu.

Sắp xếp lại số liệu mẫu:

20	100	30	980	440	20	20	150	60	270
20	20	20	30	60	100	150	270	440	980

Kích thước mẫu là $N = 10$. Số liệu thứ 5 và 6 lần lượt là 60, 100. Vậy giá trị đại diện cho bảng số liệu là $M_e = \frac{60 + 100}{2} = 80$ (Nghìn con).

Bài 4. Một cửa hàng bán 3 loại hoa quả nhập khẩu: Bưởi, Dưa vàng và Lê với số liệu tính toán được cho bởi bảng (trong một quý) sau khi giảm giá mỗi loại lần lượt là x, y, z trên 1 kg

Loại quả	Lê	Dưa vàng	Bưởi
Giá bán (nghìn/1kg)	$200 - x$	$300 - y$	$400 - z$
Số lượng bán (kg)	$200 + x$	$300 + y$	$400 + z$

Biết rằng $x + y + z = 90$ (nghìn). Tính giá trị x, y, z để lợi nhuận bình quân của một 1 kg hoa quả đạt được cao nhất.

Lời giải. Do khối lượng hoa quả bán được là $200 + x + 300 + y + 400 + z = 990$ là cố định, vì thế bình quân mỗi kg hoa quả có giá cao nhất khi tổng số tiền thu được là cao nhất.

$$\text{Tổng số tiền thu được là } P = (200 - x)(200 + x) + (300 - y)(300 + y) + (400 - z)(400 + z) \\ = 290000 - (x^2 + y^2 + z^2)$$

Ta có bất đẳng thức sau $x^2 + y^2 + z^2 \geq \frac{1}{3}(x + y + z)^2 = 2700$ từ đó $P \leq 287300$. Vậy P lớn nhất khi dấu bằng xảy ra tức là $x = y = z = 30$ (nghìn)

Bài 5. Để đảm bảo bảng số liệu được phân bố đồng đều người ta điều chỉnh các giá trị x, y sao cho số trung bình cộng và số trung vị bằng nhau. Khi đó bảng số liệu được cho như sau:

Giá trị	40	50	$60 + x$	$90 + y$	
Tần số	20	30	$60 - x$	$90 - y$	200

Biết rằng $x \leq 0$. Tìm x ?

Lời giải. Số trung bình của bảng số liệu bằng: $\bar{x} = \frac{(60 + x)(60 - x) + (90 - y)(90 + y) + 2300}{200}$ và số trung

$$\text{vị là } M_e = \frac{30 + 60 - x}{2}.$$

Ta có $20 + 30 + 60 - x + 90 - x = 200 \Leftrightarrow x + y = 0$.

Ta giải phương trình 2 giá trị này bằng nhau rồi lấy số nguyên gần nhất với nghiệm.

$$\frac{(60 + x)(60 - x) + (90 - y)(90 + y) + 2300}{200} = \frac{40 + 50 - x}{2}.$$

Do đó $x = 25 - 25\sqrt{5}$ hoặc $x = 25 + 5\sqrt{21}$, vì $x \leq 0$ nên ta lấy nghiệm thứ nhất, số nguyên gần với nó nhất là $25 - 25\sqrt{5}$.

Dạng 3. Một

Áp dụng định nghĩa về Một của bảng số liệu thống kê.

Ví dụ 5. Tuổi thọ của 30 bóng đèn được thắp thử (đơn vị: giờ) được cho bởi bảng số liệu thống kê dưới đây

1180	1150	1190	1170	1180	1170	1160	1170	1160	1150
1190	1180	1170	1170	1170	1190	1170	1170	1170	1180
1170	1160	1160	1160	1170	1160	1180	1180	1150	1170

Hãy tính một của bảng số liệu thống kê trên.

Lời giải. Từ bảng số liệu trên ta suy ra bảng phân bố tần số các giá trị tuổi thọ của 30 bóng đèn như sau

Tuổi thọ (giờ)	1150	1160	1170	1180	1190	Tổng
Tần số	3	6	12	6	3	30

Ta thấy giá trị 1170 có tần số bằng 12 là lớn nhất. Do đó mốt của bảng số liệu là: $M_O = 1170$.

Ví dụ 6. Kết quả kiểm tra chất lượng đầu năm (thang điểm 30) của 41 học sinh của một lớp được cho bởi bảng số liệu thống kê dưới đây

Điểm	9	11	14	16	17	18	20	21	23	25	Tổng
Tần số	3	7	4	4	6	7	3	3	2	2	41

Hãy tính mốt của bảng số liệu thống kê trên.

Lời giải. Ta thấy điểm 11 và điểm 18 có tần số bằng 7 là lớn nhất. Do đó bảng số liệu có hai mốt là: $M_O^{(1)} = 11$ và $M_O^{(2)} = 18$.

Ví dụ 7. Một bác sĩ mắt ghi lại tuổi của 30 bệnh nhân mắc bệnh đau mắt hột. Kết quả thu được mẫu số liệu như sau

21	17	22	18	20	17	15	13	15	20
15	12	18	17	25	17	21	15	12	18
16	23	14	18	19	13	16	19	18	17

Tính mốt M_O của bảng số liệu đã cho.

Lời giải. Từ bảng số liệu trên ta suy ra bảng phân bố tần số tuổi của 30 bệnh nhân đau mắt hột như sau

Tuổi	12	13	14	15	16	17	18	19	20	21	22	23	25	Tổng
Tần số	2	2	1	4	2	5	5	2	2	2	1	1	1	30

Ta thấy tuổi 17 và 18 có tần số bằng 5 là lớn nhất. Do đó bảng số liệu có hai mốt là: $M_O^{(1)} = 17$ và $M_O^{(2)} = 18$.

Ví dụ 8. Điểm bài kiểm tra một tiết môn toán của 40 học sinh lớp 11A1 được thống kê bằng bảng số liệu dưới đây

Điểm	3	4	5	6	7	8	9	10	Cộng
Số học sinh	2	3	$3n - 8$	$2n + 4$	3	2	4	5	40

Trong đó $n \in \mathbb{N}, n \geq 4$. Tính mốt của bảng số liệu thống kê đã cho.

Lời giải. Vì tổng các số liệu thống kê bằng 40 nên ta có: $5n + 15 = 40 \Leftrightarrow n = 5$.

Với $n = 5$ ta có bảng phân bố tần số

Điểm	3	4	5	6	7	8	9	10	Cộng
Số học sinh	2	3	7	14	3	2	4	5	40

Vậy mốt của bảng số liệu là: $M_O = 6$.

BÀI TẬP TỰ LUYỆN

Bài 6. Cho bảng phân bố tần số sau

Giá trị	x_1	x_2	x_3	x_4	x_5
Tần số	12	5	n^2	16	$6n - 5$

Tìm tất cả các số tự nhiên n để $M_O = x_3$ là một duy nhất của bảng phân bố tần số đã cho.

Lời giải. Từ giả thiết $M_O = x_3$ là một duy nhất của bảng số liệu thống kê đã cho nên ta có $\begin{cases} n^2 > 16 \\ n^2 > 6n - 5 \end{cases} \Leftrightarrow \begin{cases} n < -4 \\ n > 4 \end{cases} \cdot \Leftrightarrow \begin{cases} n < -4 \\ n > 5 \end{cases} \cdot$

Vì n là số tự nhiên nên các giá trị n thỏa mãn là: $n > 5$.

Bài 7. Cho bảng phân bố tần số sau

Giá trị	x_1	x_2	x_3	x_4	x_5
Tần số	5	2	n	$20 - n$	8

Tìm các số tự nhiên n để $M_O = x_4$ là một duy nhất của bảng số liệu thống kê đã cho.

Lời giải. Từ giả thiết $M_O = x_4$ là một duy nhất của bảng số liệu thống kê đã cho nên ta có $\begin{cases} 20 - n > 8 \\ 20 - n > n \end{cases} \Leftrightarrow \begin{cases} n < 12 \\ n < 10 \end{cases} \Leftrightarrow n < 10.$

Vì n là số tự nhiên nên các giá trị n thỏa mãn là: $0 \leq n < 10$.

Bài 8. Cho bảng phân bố tần số sau

Giá trị	x_1	x_2	x_3	x_4	x_5	x_6
Tần số	5	$n^2 + 3$	3	$7n - 9$	$n + 1$	7

Gọi S là tập hợp tất cả các số n nguyên dương sao cho $M_O = x_2$ và $M_O = x_4$ là hai một của bảng phân bố tần số đã cho. Tính số phần tử của tập hợp S .

Lời giải. Từ giả thiết x_2 và x_4 là các một của bảng số liệu thống kê đã cho, ta có

$$\begin{cases} n^2 + 3 = 7n - 9 \\ 7n - 9 > n + 1 \\ 7n - 9 > 7 \end{cases} \Leftrightarrow \begin{cases} n^2 - 7n + 12 = 0 \\ n > \frac{5}{3} \\ n > \frac{16}{7} \end{cases} \Leftrightarrow \begin{cases} n = 3 \\ n = 4 \end{cases}.$$

Vì n là số nguyên dương nên $n = 3$ và $n = 4$ thỏa mãn. Vậy tập hợp S có 2 phần tử.

BÀI TẬP TỔNG HỢP

Bài 9. Quan sát 9 con chuột chạy qua một mê hồn trận và ghi lại thời gian (tính bằng phút) của chúng trong bảng sau:

Con chuột	1	2	3	4	5	6	7	8	9
Thời gian chạy	1	2,5	3	1,5	2	1,25	1	0,9	30

a) Tính số trung bình, số trung vị và một của thời gian chuột ra khỏi mê hồn trận?

b) Trong trường hợp này nên chọn đại lượng nào để thể hiện xu thế trung bình của mẫu?

Lời giải.

a) Số trung bình: $\bar{x} = \frac{1 + 2,5 + \dots + 30}{9} \approx 4,79.$

Sắp xếp dãy số liệu theo thứ tự không giảm ta được:

$$0,9; 1; 1; 1,25; 1,5; 2; 2,5; 3; 30$$

Số trung vị: $M_e = x_5 = 1,5.$

Mốt: $M_0 = 1.$

b) Trong trường hợp này ta nên chọn số trung vị để thể hiện xu thế trung bình của mẫu.

Bài 10. Trong kỳ thi Tiếng Anh cấp chứng chỉ B1 theo chuẩn Châu Âu của trường Đại học Cần Thơ, điểm thi của 32 thí sinh (thang điểm 100) như sau:

79	65	85	52	81	55	65	49
42	68	66	56	57	65	72	69
60	50	63	74	88	78	95	41
87	61	72	53	47	90	74	68

a) Lập bảng phân bố tần số ghép lớp của mẫu số liệu trên với các lớp $[40; 50)$, $[50; 60)$, ..., $[90; 100)$.

b) Tính số trung bình của mẫu số liệu theo bảng phân bố tần số ghép lớp (chính xác đến hàng phần trăm).

c) Tính số trung vị.

Lời giải.

a) Bảng phân bố tần số ghép lớp của mẫu số liệu:

Lớp	Giá trị đại diện	Tần số
$[40; 50)$	45	4
$[50; 60)$	55	6
$[60; 70)$	65	10
$[70; 80)$	75	6
$[80; 90)$	85	4
$[90; 100)$	95	2

b) Số trung bình $\bar{x} = \frac{45 \times 4 + 55 \times 6 + \dots + 95 \times 2}{32} \approx 66,88$.

c) Vì có 32 số liệu nên số trung vị là $M_e = \frac{x_{16} + x_{17}}{2}$.

Nhìn vào bảng trên ta thấy x_{16}, x_{17} thuộc lớp $[60; 70)$, từ đó ta có $x_{16} = 65, x_{17} = 66$.

Vậy $M_e = \frac{x_{16} + x_{17}}{2} = \frac{65 + 66}{2} = 65,5$.

Bài 11. Điểm kiểm tra của Darell (thang điểm 100) trong giai đoạn đầu tiên được biểu diễn bởi dãy số liệu sau:

78, 83, 84, 86, 87, 90, 92, 92

Darell có thể miêu tả điểm của mình với bố mẹ như thế nào? Dùng số trung bình, số trung vị hay mốt? Liệu số liệu này có cho phép đưa đến một sự miêu tả đúng hay không?

Lời giải. Số trung bình: $\bar{x} = \frac{78 + 83 + 84 + 86 + 87 + 90 + 92 + 92}{8} = 86,5$.

Số trung vị: $M_e = 86,5$.

Mốt: $M_0 = 92$.

Do mốt có giá trị lớn nhất nên Darell có thể dùng nó để mô tả cho điểm kiểm tra của anh ấy.

Tuy nhiên, trong trường hợp này mốt không phải là giá trị đại diện tốt cho điểm kiểm tra của anh ấy vì tất cả các điểm còn lại đều nhỏ hơn 92.

Bài 12. Nói về tiền lương, Amara đang tham gia phỏng vấn tìm việc làm tại một công ty kỹ thuật. Cô ấy nói rằng mức lương trung bình của 37 công nhân tại đây là nhiều hơn 40000 USD. Sử dụng những thông tin có trong bảng cho biết liệu Amara có nên mong đợi mức lương hơn 40000 USD nếu cô làm việc cho công ty này. Hãy giải thích tại sao?

Người lao động	Mức lương (USD)
Giám đốc	375000
Phó giám đốc	325000
Nhân viên bán hàng (15)	35000
Thư ký (10)	16000
Nhân viên trực điện thoại (10)	12000

Lời giải. Công ty có 37 nhân viên, nhưng chỉ có hai nhân viên có mức lương trên 40000 USD là giám đốc và phó giám đốc. Do đó, Amara không nên mong đợi một mức lương cao hơn 40000 USD khi vào làm việc cho công ty này. Nói cách khác, con số trung bình về mức lương trên 40000 USD là một số không đại diện tốt cho các số liệu trên.

Bài 13. Cho bảng phân bố tần số

Giá trị	x_1	x_2	x_3	x_4	x_5
Tần số	2	$x + y$	$2x - y$	5	6

với x, y là các số tự nhiên. Tìm tất cả các cặp số $(x; y)$ để x_5 là một của bảng số liệu đã cho.

Lời giải. Điều kiện để x_5 là một của bảng số liệu đã cho là:
$$\begin{cases} 0 \leq x + y \leq 6 & (1) \\ 0 \leq 2x - y \leq 6 & (2) \end{cases}$$

Từ (1) suy ra $y \leq 6 - x$ và $x \leq 6 - y$.

Từ (2) suy ra $2x - 6 \leq y$ và $\frac{y}{2} \leq x$.

$$\text{Do đó } \begin{cases} 2x - 6 \leq 6 - x \\ \frac{y}{2} \leq 6 - y \end{cases} \Leftrightarrow \begin{cases} 0 \leq x \leq 4 \\ 0 \leq y \leq 4 \end{cases}$$

Từ đó tìm được 14 cặp số thỏa mãn là:

$(0; 0), (1; 0), (1; 1), (1; 2), (2; 0), (2; 1), (2; 2)$

$(3; 0), (3; 1), (3; 2), (3; 3), (4; 2), (2; 3), (2; 4).$

Bài 14. Cho bảng phân bố tần số

Giá trị	x_1	x_2	x_3	x_4	x_5
Tần số	6	$3x + y$	$3y - 3x$	$x + y$	4

với x, y là các số tự nhiên. Tìm tất cả các cặp số $(x; y)$ để bảng số liệu thống kê đã cho có một là 3 giá trị khác nhau.

Lời giải. Trường hợp 1: các giá trị x_1, x_2, x_3 là một khi
$$\begin{cases} 3x + y = 6 \\ 3y - 3x = 6 \\ x + y < 6 \end{cases} \Leftrightarrow \begin{cases} x = 1 \\ y = 3 \\ x + y < 6 \end{cases} \Leftrightarrow \begin{cases} x = 1 \\ y = 3 \end{cases}$$

Trường hợp 2: các giá trị x_1, x_2, x_4 là một khi
$$\begin{cases} 3x + y = 6 \\ x + y = 6 \\ 3y - 3x < 6 \end{cases} \Leftrightarrow \begin{cases} x = 0 \\ y = 6 \\ y - x < 2 \end{cases}$$

Hệ vô nghiệm vì $\begin{cases} x = 0 \\ y = 6 \end{cases}$ không thỏa mãn bất phương trình $y - x < 2$.

Trường hợp 3: các giá trị x_1, x_3, x_4 là một khi
$$\begin{cases} x + y = 6 \\ 3y - 3x = 6 \\ 3x + y < 6 \end{cases} \Leftrightarrow \begin{cases} x = 2 \\ y = 4 \\ 3x + y < 6 \end{cases}$$

Hệ vô nghiệm vì $\begin{cases} x = 2 \\ y = 4 \end{cases}$ không thỏa mãn bất phương trình $3x + y < 6$.

Trường hợp 4: các giá trị x_2, x_3, x_4 là một khi
$$\begin{cases} 3x + y = 3y - 3x \\ 3x + y = x + y \\ x + y > 6 \end{cases} \Leftrightarrow \begin{cases} x = 0 \\ y = 0 \\ x + y > 6 \end{cases}$$

Hệ vô nghiệm vì $\begin{cases} x = 0 \\ y = 0 \end{cases}$ không thỏa mãn bất phương trình $x + y > 6$.

Vậy chỉ có $\begin{cases} x = 1 \\ y = 3 \end{cases}$ thỏa mãn yêu cầu của bài toán.

§4. PHƯƠNG SAI VÀ ĐỘ LỆCH CHUẨN

I. Tóm tắt lí thuyết

Để đo độ phân tán (độ chênh lệch) giữa các giá trị của mẫu số liệu so với số trung bình, người ta đưa ra hai số đặc trưng là *phương sai* và *độ lệch chuẩn*.

Định nghĩa 1. Giả sử ta có một mẫu số liệu kích thước n là x_1, x_2, \dots, x_n .

Phương sai của mẫu số liệu này, kí hiệu là s_x^2 , được tính bởi công thức sau

$$s_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

trong đó \bar{x} là số trung bình của mẫu số liệu.

Căn bậc hai của phương sai được gọi là *độ lệch chuẩn*, kí hiệu là s_x .

$$s_x = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$

Chú ý

- Phương sai và độ lệch chuẩn càng lớn thì độ phân tán của các số liệu thống kê càng lớn.*
- Phương sai s_x^2 và độ lệch chuẩn s_x đều được dùng để đánh giá mức độ phân tán của các số liệu thống kê (so với số trung bình cộng). Nhưng khi cần chú ý đến đơn vị đo thì ta dùng s_x vì s_x có cùng đơn vị đo với dấu hiệu được nghiên cứu.*
- Phương sai còn được tính theo các công thức sau đây*

(a) *Trường hợp bảng phân bố tần số, tần suất:*

$$s_x^2 = \frac{1}{n} \sum_{i=1}^k n_i (x_i - \bar{x})^2 = \sum_{i=1}^k f_i (x_i - \bar{x})^2$$

trong đó n_i , f_i lần lượt là tần số, tần suất của giá trị x_i .

(b) *Trường hợp bảng phân bố tần số, tần suất ghép lớp:*

$$s_x^2 = \frac{1}{n} \sum_{i=1}^k n_i (c_i - \bar{x})^2 = \sum_{i=1}^k f_i (c_i - \bar{x})^2$$

trong đó c_i , n_i , f_i lần lượt là giá trị đại diện, tần số, tần suất của giá trị x_i .

(c) *Người ta còn chứng minh được công thức sau:*

$$s_x^2 = \overline{x^2} - \bar{x}^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \frac{1}{n^2} \left(\sum_{i=1}^n x_i \right)^2$$

II. Các dạng toán

Dạng 1. Tính phương sai và độ lệch chuẩn của bảng số liệu KHÔNG ghép lớp

a) Để tính phương sai s^2 của một mẫu số liệu $\{x_1; x_2; \dots; x_N\}$ ta thực hiện một trong các cách sau:

• **Cách 1:**

- + Tính số trung bình: $\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$;
- + Tính các độ lệch: $x_i - \bar{x}$, ($i = \overline{1, N}$);
- + Tính các phương sai theo công thức: $s^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2$.

• **Cách 2:**

- + Tính $\sum_{i=1}^N x_i$ và $\sum_{i=1}^N x_i^2$;
- + Tính phương sai theo công thức: $s^2 = \frac{1}{N} \sum_{i=1}^N x_i^2 - \frac{1}{N^2} \left(\sum_{i=1}^N x_i \right)^2$.

Chú ý: Nếu bảng số liệu được cho bởi bảng phân phối tần số như sau:

X	x_1	x_2	...	x_m	
Tần số	n_1	n_2	...	n_m	N

Thì phương sai được tính theo công thức: $s^2 = \frac{1}{N} \sum_{i=1}^N n_i (x_i - \bar{x})^2$.

b) Tính độ lệch chuẩn s : Độ lệch chuẩn s bằng căn bậc hai của phương sai:

$$s = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}.$$

Ví dụ 1. Sản lượng lúa (đơn vị là tạ) của 40 thửa ruộng thí nghiệm có cùng diện tích được trình bày trong bảng tần số dưới đây:

Sản lượng (x)	20	21	22	23	24	
Tần số (n)	5	8	11	10	6	$N = 40$

- a) Tính sản lượng trung bình của 40 thửa ruộng?
- b) Tính phương sai và độ lệch chuẩn.

Lời giải. a) Số trung bình của sản lượng của 40 thửa ruộng là:

$$\bar{x} = \frac{5 \cdot 20 + 8 \cdot 21 + 11 \cdot 22 + 10 \cdot 23 + 6 \cdot 24}{40} = 22,1 \text{ tạ}.$$

b) Tính phương sai:

Cách 1: $s^2 = \frac{1}{N} \sum_{i=1}^5 n_i (x_i - \bar{x})^2$, thay số vào ta được:

$$s^2 = \frac{1}{40} [5(20 - 22,1)^2 + 8(21 - 22,1)^2 + 11(22 - 22,1)^2 + 10(23 - 22,1)^2 + 6(24 - 22,1)^2] = \frac{6160}{4000}.$$

Hay $s^2 = 1,54$.

Cách 2: Ta có:

- $\sum_{i=1}^5 n_i x_i = 5 \cdot 20 + 8 \cdot 21 + 11 \cdot 22 + 10 \cdot 23 + 6 \cdot 24 = 884.$
- $\sum_{i=1}^5 n_i x_i^2 = 5 \cdot 20^2 + 8 \cdot 21^2 + 11 \cdot 22^2 + 10 \cdot 23^2 + 6 \cdot 24^2 = 19598.$

Do đó $s^2 = \frac{1}{N} \sum_{i=1}^N n_i x_i^2 - \frac{1}{N} \left(\sum_{i=1}^N n_i x_i \right)^2 = \frac{1}{40} \cdot 19598 - \frac{1}{40^2} \cdot 884^2 = 1,54.$

Tính độ lệch chuẩn: $s = \sqrt{s^2} = \sqrt{1,54} \approx 1,24.$

Ví dụ 2. 100 học sinh tham gia kì thi học sinh giỏi toán (thang điểm là 20). Kết quả được cho trong bảng sau:

Điểm	9	10	11	12	13	14	15	16	17	18	19	
Tần số	1	1	3	5	8	13	19	24	14	10	2	$N = 100$

- a) Tính sản lượng trung bình.
b) Tính phương sai và độ lệch chuẩn.

Lời giải. a) Tính số trung bình:

$$\sum_{i=1}^{11} n_i x_i = 1 \cdot 9 + 1 \cdot 10 + \dots + 10 \cdot 18 + 2 \cdot 19 = 1523.$$

Nên số trung bình là $\bar{x} = \frac{1523}{100} = 15,23.$

b) Ta có: $\sum_{i=1}^{11} n_i x_i = 1523$ và $\sum_{i=1}^{11} n_i x_i^2 = 23591$ nên phương sai là:

$$s^2 = \frac{1}{N} \sum_{i=1}^N n_i x_i^2 - \frac{1}{N^2} \left(\sum_{i=1}^N n_i x_i \right)^2 = \frac{1}{100} \cdot 23591 - \frac{1}{100^2} \cdot (1523)^2 \approx 3,96$$

Độ lệch chuẩn: $s = \sqrt{s^2} \approx 1,99.$

Ví dụ 3. Số máy tính bán được trong 7 tháng liên tiếp của một cửa hàng được ghi lại trong bảng sau:

83	79	92	71	69	83	74
----	----	----	----	----	----	----

Tính số trung bình, phương sai và độ lệch chuẩn.

Lời giải. Số trung bình là $\bar{x} = \frac{83 + 79 + 92 + 71 + 69 + 83 + 74}{7} \approx 78,71.$

Ta có $\frac{1}{N} \sum_{i=1}^N x_i^2 = 6251,57, \frac{1}{N^2} \left(\sum_{i=1}^N x_i \right)^2 = 6195,94.$ Suy ra

$$s^2 = \frac{1}{N} \sum_{i=1}^N x_i^2 - \frac{1}{N^2} \left(\sum_{i=1}^N x_i \right)^2 = 6251,57 - 6195,94 = 55,63.$$

Vậy $s = \sqrt{55,63} \approx 7,46.$

BÀI TẬP TỰ LUYỆN

Bài 1. Kết quả thi kết thúc học kì một của bạn Hoa được ghi lại trong bảng sau:

Văn	Địa	Lý	Hóa	Toán	Anh văn
6,0	8,0	7,5	8,5	7,0	7,5

Tìm số trung bình, phương sai và độ lệch chuẩn.

Lời giải. $\bar{x} = 7,5; s^2 \approx 0,42; s \approx 0,65$.

Bài 2. Theo dõi số áo bán ra của 9 loại áo tại một cửa hàng, người ta có dãy số liệu sau (đơn vị: chiếc)

42	52	23	36	48	42	40	48	42
----	----	----	----	----	----	----	----	----

Tìm số trung bình, phương sai và độ lệch chuẩn.

Lời giải. $\bar{x} = 41,1; s^2 \approx 63,4; s \approx 8,0$.

Bài 3. Trong sổ theo dõi bán hàng ở một cửa hàng bán xe máy có bảng sau:

Số xe bán trong ngày	0	1	2	3	4	5
Tần số	2	13	15	12	7	3

Tìm số xe trung bình bán được trong ngày. Tìm phương sai và độ lệch chuẩn.

Lời giải. Trung bình cộng: 48,35; phương sai $s^2 \approx 194,64$; độ lệch chuẩn 13,95.

Bài 4. Bảng số liệu sau cho ta lãi (quy tròn) hàng tháng của một cửa hàng A trong năm 2006 (đơn vị là triệu đồng).

Tháng	1	2	3	4	5	6	7	8	9	10	11	12
Lãi	12	15	18	13	18	16	17	14	18	17	20	17

Tìm số trung bình. Tìm phương sai và độ lệch chuẩn.

Lời giải. Trung bình cộng: 16,25; phương sai $s^2 \approx 5,02$; độ lệch chuẩn 2,24.

Bài 5. Theo dõi số bao xi măng bán ra trong 22 ngày tại một cửa hàng bán vật liệu xây dựng ta có bảng sau:

47	54	43	50	61	36	65	54	43	50	62
59	36	45	45	33	53	67	21	45	50	36

Tìm số trung bình. Tìm phương sai và độ lệch chuẩn.

Lời giải. Trung bình cộng: 47,95; phương sai $s^2 \approx 123,13$; độ lệch chuẩn 11,09.

Bài 6. Bảng sau đây ghi lại tốc độ (km/h) của 30 chiếc ô tô.

60	65	70	68	62	75	80	83	82	69
73	75	85	72	67	88	90	85	72	63
75	76	85	84	70	61	60	65	73	76

Tìm số trung bình. Tìm phương sai và độ lệch chuẩn.

Lời giải. Trung bình cộng: 70,70; phương sai $s^2 \approx 38,21$; độ lệch chuẩn 6,18.

Bài 7. Số liệu sau đây cho ta số lãi mỗi tháng của một cửa hàng năm 2004 (đơn vị: triệu đồng).

Tháng	1	2	3	4	5	6	7	8	9	10	11	12
Lãi	12	15	18	13	13	16	18	19	15	17	20	17

Tìm số trung bình. Tìm phương sai và độ lệch chuẩn.

Lời giải. Trung bình cộng: 16; phương sai $s^2 \approx 5,9$; độ lệch chuẩn 2,43.

Dạng 2. Tính phương sai và độ lệch chuẩn của bảng số liệu ghép lớp

Để tính phương sai của bảng phân bố tần số, tần suất ghép lớp ta dùng công thức

$$s_x^2 = \frac{1}{n} \left[n_1 (c_1 - \bar{x})^2 + n_2 (c_2 - \bar{x})^2 + \cdots + n_k (c_k - \bar{x})^2 \right]$$

trong đó c_i, n_i, f_i lần lượt là giá trị đại diện, tần số, tần suất của lớp thứ i ; c_i được tính bằng trung bình cộng của 2 giá trị đầu mút của lớp i , n là số các số liệu thống kê ($n = n_1 + n_2 + \cdots + n_k$), \bar{x} là số trung bình cộng của các số trong số liệu thống kê đã cho.

⚠ Người ta còn chứng minh được công thức $s_x^2 = \overline{x^2} - (\bar{x})^2$.

Độ lệch chuẩn s_x được tính bởi công thức $s_x = \sqrt{s_x^2}$.

Ví dụ sau sử dụng công thức $s_x^2 = \frac{1}{n} [n_1(x_1 - \bar{x})^2 + \dots + n_k(x_k - \bar{x})^2]$ để tính phương sai.

Ví dụ 4. Cho bảng phân bố tần số ghép lớp sau

Độ dài của 60 lá dương xỉ trưởng thành

Lớp của độ dài (cm)	Tần số
[10; 20)	8
[20; 30)	18
[30; 40)	24
[40; 50]	10
Cộng	60

Tính phương sai và độ lệch chuẩn của bảng phân bố tần số ghép lớp đã cho.

Lời giải. Trước hết ta có $\bar{x} = \frac{15 \times 8 + 25 \times 18 + 35 \times 24 + 45 \times 10}{60} = 31$.

Khi đó phương sai $s_x^2 = \frac{8(15 - 31)^2 + 18(25 - 31)^2 + 24(35 - 31)^2 + 10(45 - 31)^2}{60} = 84$.

Độ lệch chuẩn $s_x = \sqrt{84} \approx 9,17$.

Ví dụ sau sử dụng công thức $s_x^2 = \overline{x^2} - (\bar{x})^2$ để tính phương sai.

Ví dụ 5. Cho bảng phân bố tần số ghép lớp sau

Khối lượng của 30 củ khoai tây

Lớp của khối lượng (g)	Tần số
[70; 80)	3
[80; 90)	6
[90; 100)	12
[100; 110)	6
[110; 120)	3
Cộng	30

Tính phương sai và độ lệch chuẩn của bảng phân bố tần số ghép lớp đã cho.

Lời giải. Ta tính các giá trị $n_i c_i$ và $n_i c_i^2$ và bổ sung vào bảng đã cho, ta được bảng sau

Lớp của khối lượng (g)	Tần số	$n_i c_i$	$n_i c_i^2$
[70; 80)	3	225	16875
[80; 90)	6	510	43350
[90; 100)	12	1140	108300
[100; 110)	6	630	66150
[110; 120)	3	345	39675
Cộng	30	2850	274350

Từ đó, ta tính được $\bar{x} = 95$ và $\overline{x^2} = 9145$. Áp dụng công thức $s_x^2 = \overline{x^2} - (\bar{x})^2$, ta tính được $s_x^2 = 120$ và $s_x = \sqrt{s_x^2} \approx 10,95$.

Ví dụ sau cho bảng phân bố tần suất ghép lớp. Ta tính \bar{x} và s_x^2 dựa trên tần suất.

Ví dụ 6. Cho bảng phân bố tần suất ghép lớp sau

Chiều cao của 35 cây bạch đàn

Lớp của chiều cao (m)	Tần suất (%)
[6,5;7,0)	5,7
[7,0;7,5)	11,4
[7,5;8,0)	25,7
[8,0;8,5)	31,4
[8,5;9,0)	17,2
[9,0;9,5]	8,6
Cộng	100

Tính phương sai và độ lệch chuẩn của bảng phân bố tần suất ghép lớp đã cho.

Lời giải. Trước hết ta tính ra các giá trị $f_i c_i$, cuối bảng sẽ có được \bar{x} , từ đó tính $f_i(c_i - \bar{x})^2$, cuối bảng sẽ có s_x^2 .

Lớp của chiều cao (m)	Tần suất (%)	$f_i c_i$	$f_i(c_i - \bar{x})^2$
[6,5;7,0)	5,7	0,38475	0,102961
[7,0;7,5)	11,4	0,8265	0,081206
[7,5;8,0)	25,7	1,99175	0,030412
[8,0;8,5)	31,4	2,5905	0,007642
[8,5;9,0)	17,2	1,505	0,074018
[9,0;9,5]	8,6	0,7955	0,114925
Cộng	100	8,094	0,411164

Như vậy ta được phương sai $s_x^2 = 0,411164$, suy ra $s_x \approx 0,641221$.

Ví dụ sau sử dụng sự hỗ trợ của máy tính *fx-570ESPLUS* để tính phương sai và độ lệch chuẩn.

Ví dụ 7. Cho bảng phân bố tần số ghép lớp sau

Khối lượng của nhóm 20 cá mè

Lớp khối lượng (kg)	[0,6;0,8)	[0,8;1,0)	[1,0;1,2)	[1,2,1,4]	Cộng
Tần số	4	6	6	4	20

Tính phương sai và độ lệch chuẩn của bảng trên.

Lời giải. Trước hết, ta chọn **[shift] [mode] ▾ 4 1]** để dùng chế độ thống kê với 1 đối tượng thống kê.

Sau đó ta vào **[mode] 3]** để vào chế độ thống kê và chọn **1]** để nhập dữ liệu.

Nhập các giá trị đại diện trong cột X trên màn hình. Sau khi nhập xong, chuyển qua cột FREQ bằng phím **[>]** và nhập các tần số tương ứng với các giá trị đại diện.

Nhập xong bấm **[AC]**. Để tính độ lệch chuẩn, ta bấm **[shift] 1 4 3 [=]**, kết quả là $s_x = 0,2049390153$, ta tính phương sai bằng cách bình phương giá trị trên, bấm tiếp **[x²] [=]**, ta được $s_x^2 = 0,042$.

BÀI TẬP TỰ LUYỆN

Bài 8. Nhiệt độ trung bình của tháng 2 ở một thành phố đo trong 30 năm được cho trong bảng sau.

Lớp nhiệt độ (°C)	Tần số	Tần suất (%)
[12;14)	1	3,33
[14;16)	3	10,00
[16;18)	12	40,00
[18;20)	9	30,00
[20;22]	5	16,67
Cộng	30	100

Tính phương sai và độ lệch chuẩn của bảng phân bố đã cho.

Lời giải. Sử dụng một trong các phương án tính như trong ví dụ, ta được kết quả: $s_x^2 \approx 3,93$; $s_x \approx 1,98$.

Bài 9. Tính phương sai và độ lệch chuẩn của bảng phân bố sau

Chiều cao của 36 học sinh

Lớp số đo chiều cao (cm)	Tần số
[150; 156)	6
[156; 162)	12
[162; 168)	13
[168; 174]	5
Cộng	36

Lời giải. Sử dụng một trong các phương án tính như trong ví dụ, ta được kết quả: $s_x^2 \approx 30,97$; $s_x = 5,57$.

Bài 10. Tính phương sai và độ lệch chuẩn của bảng phân bố sau

Tiền lãi của mỗi ngày bán báo được khảo sát trong 30 ngày

Lớp tiền lãi (nghìn đồng)	Tần số
[29, 5; 40, 5)	3
[40, 5; 51, 5)	4
[51, 5; 62, 5)	9
[62, 5; 73, 5)	5
[73, 5; 84, 5)	5
[84, 5; 95, 5]	4
Cộng	30

Lời giải. Sử dụng một trong các phương án tính như trong ví dụ, ta được kết quả: $s_x^2 = 271,71$; $s_x = 16,48$.

Bài 11. Trong một trường THPT, cho kiểm tra toán ở 2 lớp 10A và 10B và lập được bảng tần số ghép lớp như sau:

Điểm thi toán của lớp 10A

Lớp điểm kiểm tra	Tần số
[0; 2)	2
[2; 4)	4
[4; 6)	12
[6; 8)	28
[8; 10]	4
Cộng	50

Điểm thi toán của lớp 10B

Lớp điểm kiểm tra	Tần số
[0; 2)	4
[2; 4)	10
[4; 6)	18
[6; 8)	14
[8; 10]	5
Cộng	50

Tính phương sai và độ lệch chuẩn của hai bảng phân bố tần số ghép lớp trên và cho kết luận.

Lời giải. Sử dụng một trong các phương án tính như trong ví dụ, ta được kết quả

- Lớp 10A: $s_x^2 = 3,23$; $s_x = 1,8$.
- Lớp 10B: $s_x^2 = 4,65$; $s_x = 2,16$.

Từ đó cho thấy độ phân tán của lớp 10B nhiều hơn độ phân tán của lớp 10A so với giá trị trung bình của dữ liệu.

Bài 12. Một nông dân nuôi cá có 2 ao nuôi cùng một loại cá. Ông ta bắt mỗi ao 24 con cá và cân. Sau đây là bảng phân bố khối lượng 2 nhóm cá.

<i>Nhóm cá thứ nhất</i>	
Lớp khối lượng (g)	Tần số
[630; 635)	1
[635; 640)	2
[640; 645)	3
[645; 650)	6
[650; 655]	12
Cộng	

<i>Nhóm cá thứ hai</i>	
Lớp khối lượng (g)	Tần số
[630; 635)	0
[635; 640)	0
[640; 645)	8
[645; 650)	7
[650; 655]	9
Cộng	

Tính phương sai và độ lệch chuẩn của hai bảng phân bố tần số ghép lớp trên và cho kết luận.

Lời giải. Sử dụng một trong các phương án tính như trong ví dụ, ta được kết quả

- Nhóm cá thứ nhất: $s_x^2 = 33,16$; $s_x = 5,76$.
- Nhóm cá thứ hai: $s_x^2 = 17,66$; $s_x = 4,2$.

Từ đó cho thấy độ phân tán của nhóm cá thứ hai ít hơn độ phân tán của nhóm cá thứ nhất so với giá trị trung bình của dữ liệu.

BÀI TẬP TỔNG HỢP

Bài 13. Một trang trại trồng hai loại táo A và B. Chủ trang trại phải lựa chọn một loại táo có trọng lượng các quả táo ít bị phân tán để xuất khẩu. Sau vụ thu hoạch, ông đã cân trọng lượng của 100 quả táo. Các số liệu được tóm tắt trong bảng tần số sau:

Trọng lượng các quả táo loại A

Trọng lượng	Tần số
[80; 120)	9
[120; 160)	13
[160; 200)	5
[200; 240)	15
[240; 280)	8
Tổng	50

Trọng lượng các quả táo loại B

Trọng lượng	Tần số
[80; 120)	8
[120; 160)	11
[160; 200)	12
[200; 240)	11
[240; 280)	8
Tổng	50

Em hãy cho biết chủ trang trại sẽ chọn loại táo nào để xuất khẩu?

Lời giải. Ta có $\bar{x}_A = \bar{x}_B = 180$.

Áp dụng công thức $s_x^2 = \frac{1}{n} \sum_{i=1}^k n_i (c_i - \bar{x})^2$ ta được phương sai của mỗi bảng số liệu lần lượt là

$$s_A^2 = 3072$$

$$s_B^2 = 2752$$

Vì $s_B^2 < s_A^2$ nên trọng lượng các quả táo loại B sẽ đồng đều hơn loại A . Do đó chủ trang trại sẽ chọn loại táo B để xuất khẩu.

Bài 14. Trên hai con đường A và B , trạm kiểm soát đã ghi lại tốc độ (km/h) của 30 chiếc ô tô trên mỗi con đường như sau:

Con đường A :

60	65	70	68	62	75	80	83	82	69	73	75	85	72	67
88	90	85	72	63	75	76	85	84	70	61	60	65	73	76

Con đường B :

76	64	58	82	72	70	68	75	63	67	74	70	79	74	60
80	73	75	71	68	72	73	79	80	63	62	71	70	69	63

- Tính số trung bình, phương sai và độ lệch chuẩn của tốc độ ô tô trên mỗi con đường A , B .
- Theo em thì chạy xe trên con đường nào an toàn hơn?

Lời giải.

- Trên con đường A . Ta có: $\bar{x}_A \approx 73,63$ km/h, $s_A^2 \approx 74,77$, $s_A \approx 8,65$ km/h.
Trên con đường B . Ta có: $\bar{x}_B \approx 70,7$ km/h, $s_B^2 \approx 38,21$, $s_B \approx 6,18$ km/h.

- Nhận xét: Trên con đường B , tốc độ trung bình và độ lệch chuẩn đều nhỏ hơn trên con đường A .
Do đó chạy xe trên con đường B sẽ an toàn hơn trên con đường A .

§5. ĐỀ KIỂM TRA CHƯƠNG V

I. Đề số 1a

Bài 1. (3,5 điểm) Hai lớp 10A và 10B của một trường THPT cùng làm bài thi môn Toán, chung một đề thi. Kết quả thi được trình bày ở hai bảng phân bố tần số sau đây

Điểm	3	5	6	7	8	9	10	Cộng
Lớp 10A	7	9	3	3	7	12	4	45

Điểm	4	5	6	7	8	9	10	Cộng
Lớp 10B	6	6	7	8	9	5	4	45

- a) Hãy tính số trung bình, phương sai, độ lệch chuẩn từ các bảng phân bố tần số đã cho (làm tròn đến chữ số thập phân thứ hai).
- b) Xét xem kết quả bài thi môn Toán của lớp nào đồng đều hơn?

Lời giải.

a) **Lớp 10A:**

$$\text{Số trung bình } \bar{x}_A = \frac{3 \cdot 7 + 5 \cdot 9 + \dots + 10 \cdot 4}{45} \approx 6,87. \dots\dots\dots 0,5 \text{ điểm.}$$

$$\text{Phương sai } s_A^2 = \frac{(3 - \bar{x}_A)^2 \cdot 7 + \dots + (10 - \bar{x}_A)^2 \cdot 4}{45} \approx 5,38. \dots\dots\dots 0,5 \text{ điểm.}$$

$$\text{Độ lệch chuẩn } s_A = \sqrt{s_A^2} \approx 2,32. \dots\dots\dots 0,5 \text{ điểm.}$$

Lớp 10B:

$$\text{Số trung bình } \bar{x}_B \approx 6,87. \dots\dots\dots 0,5 \text{ điểm.}$$

$$\text{Phương sai } s_B^2 \approx 3,69. \dots\dots\dots 0,5 \text{ điểm.}$$

$$\text{Độ lệch chuẩn } s_B \approx 1,92. \dots\dots\dots 0,5 \text{ điểm.}$$

- b) Kết quả bài thi môn Toán của lớp 10B đồng đều hơn vì $s_B < s_A$. $\dots\dots\dots 0,5 \text{ điểm.}$

Bài 2. (5,0 điểm) Đo chiều cao của 40 học sinh trường THPT X, ta có bảng số liệu sau

150	151	151	151	152	152	153	153	154	155
155	156	156	156	157	159	159	160	160	161
161	162	164	165	166	166	167	167	167	168
168	169	170	170	171	171	172	173	174	175

- a) Hãy lập bảng phân bố tần số ghép lớp với các lớp là

$$[150; 154), [155; 160), [160; 165), [165; 170), [170; 175].$$

Từ đó, hãy tính số trung bình, phương sai và độ lệch chuẩn của bảng phân bố tần số ghép lớp đã lập (tính chính xác đến chữ số thập phân thứ hai).

- b) Hãy vẽ biểu đồ tần số hình cột về chiều cao của 40 học sinh trường X.

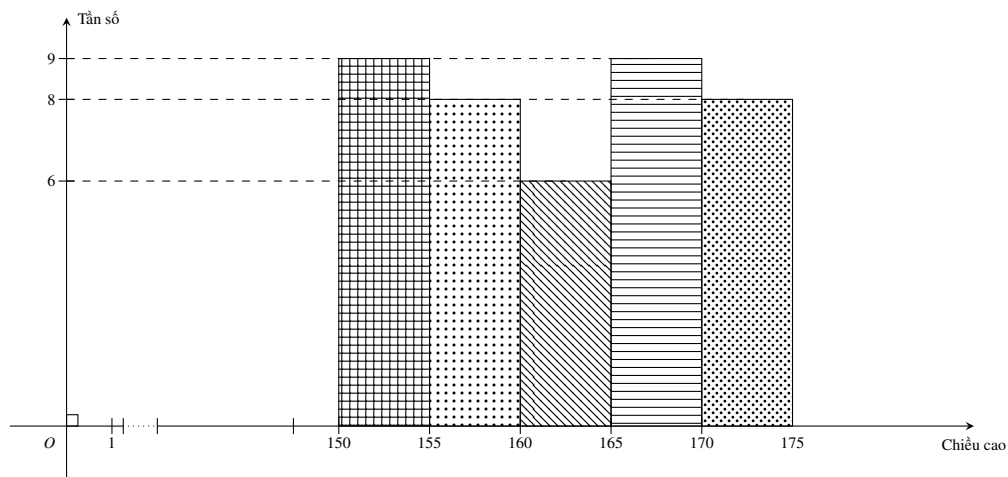
Lời giải.

- a) Bảng phân bố tần số ghép lớp

Lớp	Tần số	Giá trị đại diện
[150; 155)	9	152,5
[155; 160)	8	157,5
[160; 165)	6	162,5
[165; 170)	9	167,5
[170; 175]	8	172,5
	$n = 40$	

Lập được cột tần số 0,5 điểm.
 Lập được cột giá trị đại diện 0,5 điểm.
 Tính được giá trị trung bình $\bar{x} = 162,37$ 0,75 điểm.
 Tính được phương sai $s_x^2 \approx 53,14$ 0,75 điểm.
 Tính được độ lệch chuẩn $s_x \approx 7,29$ 0,5 điểm.

b) Biểu đồ tần số hình cột về chiều cao của 40 học sinh trường X:



Bài 3. (1,5 điểm) Một học sinh ghi lại bảng phân bố tần số, tần suất ghép lớp của một mẫu số liệu như sau

Lớp	[1; 9]	[10; 19]	[20; 29]	[30; 39]	[40; 49]	
Tần số						$n =$
Tần suất (%)	12,5	0,0	50,0	25,0	12,5	100

Tuy nhiên, em đó quên ghi kích thước mẫu n . Biết rằng n là số có 2 chữ số và chữ số tận cùng là 2. Tìm giá trị nhỏ nhất của n .

Lời giải. Hai lớp [1; 9] và [40; 49] có tần số là $n \cdot 12,5\% = \frac{n}{8}$ 0,25 điểm.

Lớp [20; 29] có tần số là $n \cdot 50\% = \frac{n}{2}$ 0,25 điểm.

Lớp [30; 39] có tần số là $n \cdot 25\% = \frac{n}{4}$ 0,25 điểm.

Vì tần số là các số nguyên dương nên n phải chia hết cho 8; 4; 2. 0,25 điểm.

Mà n là số có 2 chữ số, chữ số tận cùng là 2 và nhỏ nhất nên $n = 32$ 0,5 điểm.

II. Đề số 1b

Bài 1. (3,5 điểm) Hai lớp 10C và 10D của một trường THPT cùng làm bài thi môn Văn, chung một đề thi. Kết quả thi được trình bày ở hai bảng phân bố tần số sau đây

Điểm	10	9	8	7	6	3	3	Cộng
Lớp 10C	4	12	7	3	3	9	7	45

Điểm	10	9	8	7	6	5	4	Cộng
Lớp 10D	4	5	9	8	7	6	6	45

- a) Hãy tính số trung bình, phương sai, độ lệch chuẩn từ các bảng phân bố tần số đã cho (làm tròn đến chữ số thập phân thứ hai).
- b) Xét xem kết quả bài thi môn Toán của lớp nào đồng đều hơn?

Lời giải.**a) Lớp 10C:**

$$\text{Số trung bình } \bar{x}_C = \frac{3 \cdot 7 + 5 \cdot 9 + \dots + 10 \cdot 4}{45} \approx 6,87. \dots\dots\dots 0,5 \text{ điểm.}$$

$$\text{Phương sai } s_C^2 = \frac{(3 - \bar{x}_C)^2 \cdot 7 + \dots + (10 - \bar{x}_C)^2 \cdot 4}{45} \approx 5,38. \dots\dots\dots 0,5 \text{ điểm.}$$

$$\text{Độ lệch chuẩn } s_C = \sqrt{s_C^2} \approx 2,32. \dots\dots\dots 0,5 \text{ điểm.}$$

Lớp 10D:

$$\text{Số trung bình } \bar{x}_D \approx 6,87. \dots\dots\dots 0,5 \text{ điểm.}$$

$$\text{Phương sai } s_D^2 \approx 3,69. \dots\dots\dots 0,5 \text{ điểm.}$$

$$\text{Độ lệch chuẩn } s_D \approx 1,92. \dots\dots\dots 0,5 \text{ điểm.}$$

- b) Kết quả bài thi môn Toán của lớp 10B đồng đều hơn vì $s_D < s_C$. $\dots\dots\dots 0,5 \text{ điểm.}$

Bài 2. (5,0 điểm) Đo chiều cao của 40 học sinh trường THPT Y, ta có bảng số liệu sau

150	150	150	151	152	152	153	154	154	155
155	156	156	157	157	158	159	160	160	161
162	163	164	165	166	167	167	167	168	168
169	169	170	171	171	172	173	174	175	175

- a) Hãy lập bảng phân bố tần số ghép lớp với các lớp là

$$[150; 154), [155; 160), [160; 165), [165; 170), [170; 175].$$

Từ đó, hãy tính số trung bình, phương sai và độ lệch chuẩn của bảng phân bố tần số ghép lớp đã lập (tính chính xác đến chữ số thập phân thứ hai).

- b) Hãy vẽ biểu đồ tần số hình cột về chiều cao của 40 học sinh trường Y.

Lời giải.

- a) Bảng phân bố tần số ghép lớp

Lớp	Tần số	Giá trị đại diện
[150; 155)	9	152,5
[155; 160)	8	157,5
[160; 165)	6	162,5
[165; 170)	9	167,5
[170; 175]	8	172,5
	$n = 40$	

$$\text{Lập được cột tần số } \dots\dots\dots 0,5 \text{ điểm.}$$

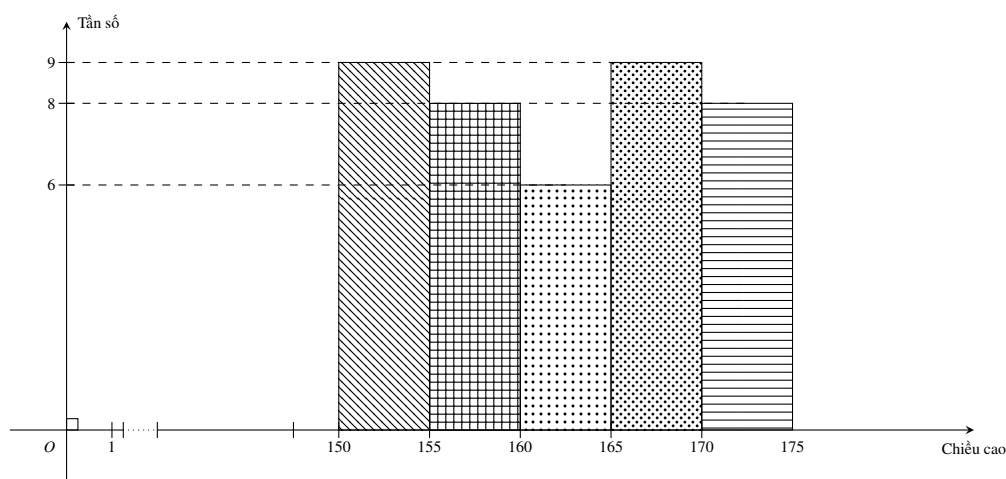
$$\text{Lập được cột giá trị đại diện } \dots\dots\dots 0,5 \text{ điểm.}$$

$$\text{Tính được giá trị trung bình } \bar{x} = 162,37. \dots\dots\dots 0,75 \text{ điểm.}$$

$$\text{Tính được phương sai } s_x^2 \approx 53,14. \dots\dots\dots 0,75 \text{ điểm.}$$

$$\text{Tính được độ lệch chuẩn } s_x \approx 7,29. \dots\dots\dots 0,5 \text{ điểm.}$$

b) Biểu đồ tần số hình cột về chiều cao của 40 học sinh trường Y:



Bài 3. (1,5 điểm) Một học sinh ghi lại bảng phân bố tần số, tần suất của một mẫu số liệu như sau

Giá trị	0	1	2	3	4	
Tần số						$n =$
Tần suất (%)	6,25	50,0	25,0	6,25	12,5	100

Tuy nhiên, em đó quên ghi kích thước mẫu n . Biết rằng n là số có 2 chữ số và chữ số tận cùng là 4. Tìm giá trị n .

Lời giải. Giá trị 0 và 3 có tần số là $n \cdot 6,25\% = \frac{n}{16}$ 0,25 điểm.

Giá trị 1 có tần số là $n \cdot 50\% = \frac{n}{2}$ 0,25 điểm.

Giá trị 2 có tần số là $n \cdot 25\% = \frac{n}{4}$ 0,25 điểm.

Giá trị 4 có tần số là $n \cdot 12,5\% = \frac{n}{8}$ 0,25 điểm.

Vì tần số là các số nguyên dương nên n phải chia hết cho 16; 8; 4; 2. 0,25 điểm.

Mà n là số có 2 chữ số, chữ số tận cùng là 4 nên $n = 64$ 0,25 điểm.

III. Đề số 2a

Bài 1. Điểm kiểm tra môn Toán cuối học kỳ 1 của một nhóm 16 học sinh là:

8, 9, 7, 0, 2, 7, 9, 10, 4, 5, 7, 8, 8, 9, 10, 7.

Tìm mốt và số trung vị của các số liệu thống kê đã cho.

Lời giải. Sắp xếp số liệu theo thứ tự không giảm ta được:

0, 2, 4, 5, 7, 7, 7, 7, 8, 8, 8, 9, 9, 9, 10, 10.

Suy ra,

• Phần tử có tần số lớn nhất là 7. Do đó mốt $M_O = 7$ **1,0 điểm.**

• Trung vị $M_e = \frac{7+8}{2} = 7,5$ **1,0 điểm**

Bài 2. Tiền thưởng của 35 nhân viên trong một công ti được thống kê trong bảng tần số ghép lớp sau đây (đơn vị: triệu đồng)

Lớp	[20; 24]	[25; 29]	[30; 34]	[35; 39]	[40; 44]	Cộng
Tần số	2	7	15	8	3	$n = 35$

Tính số trung bình và độ lệch chuẩn (kết quả làm tròn đến hàng phần trăm).

Lời giải.

• Số trung bình $\bar{x} = \frac{2 \cdot 22 + 7 \cdot 27 + 15 \cdot 32 + 8 \cdot 37 + 3 \cdot 42}{35} = 32,43$ **1,0 điểm.**

• Phương sai

$$s_x^2 = \frac{2 \cdot 22^2 + 7 \cdot 27^2 + 15 \cdot 32^2 + 8 \cdot 37^2 + 3 \cdot 42^2}{35} - \left(\frac{2 \cdot 22 + 7 \cdot 27 + 15 \cdot 32 + 8 \cdot 37 + 3 \cdot 42}{35} \right)^2.$$

Do vậy độ lệch chuẩn $s_x = \sqrt{s^2} = 4,98$ **1,0 điểm**

Bài 3. Điều tra về chiều cao của học sinh (đơn vị: cm) trong trường THPT, người ta thu được bảng số liệu thống kê sau:

160	163	166	162	164	167	165	169	163	167	171	168
165	161	170	171	164	172	162	167	173	169	164	166
164	166	165	167	161	168	164	174	165	168	163	161

a) Lập bảng phân bố tần số ghép lớp, sử dụng các lớp sau:

$$[160; 163), [163; 166), [166; 169), [169; 172), [172; 175).$$

b) Vẽ biểu đồ tần số hình cột.

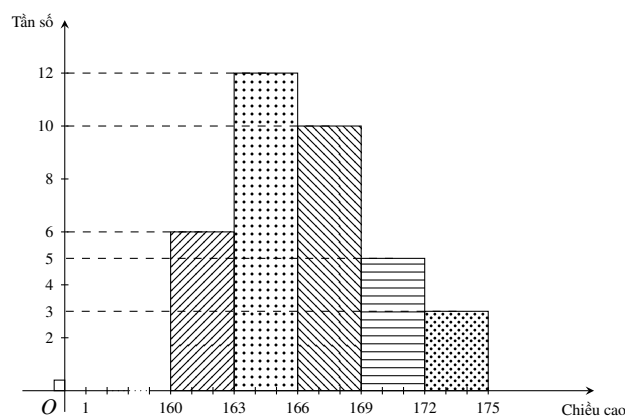
Lời giải.

a) Bảng phân bố tần số ghép lớp:

Lớp	[160; 163)	[163; 166)	[166; 169)	[169; 172)	[172; 175)	Cộng
Tần số	6	12	10	5	3	$n = 36$

..... **1,0 điểm**

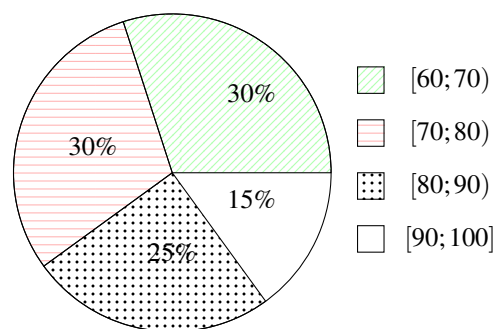
b) Biểu đồ tần số hình cột:



..... **2,0 điểm**

Bài 4.

Biểu đồ tần suất hình quạt hình bên mô tả bảng phân bố tần suất ghép lớp của dữ liệu điểm thi của 20 học sinh trong kì thi Tiếng Anh (thang điểm 100). Hãy lập bảng phân bố tần số và tần suất ghép lớp của dữ liệu đó.



Lời giải. Bảng phân bố tần số và tần suất ghép lớp:

Lớp	[60; 70)	[70; 80)	[80; 90)	[90; 100]	Cộng
Tần số	6	6	5	3	$n = 20$
Tần suất (%)	30	30	25	15	100 (%)

..... **2,0 điểm**

Bài 5. Điểm kiểm tra môn Toán của 35 học sinh lớp 10A được thống kê trong bảng phân bố tần số sau đây (thang điểm 10):

Điểm	0	1	2	3	4	5	6	7	8	9	10	Cộng
Tần số	2	1	2	1	2	3	x	5	y	4	3	$n = 35$

Biết rằng mẫu số liệu trên có 2 một. Hãy tìm x và y .

Lời giải. Tổng số học sinh là 35 nên $x + y = 12$, suy ra có ít nhất một trong hai số x hoặc y không nhỏ hơn 6. Vì mẫu số liệu có 2 một nên $x = y = 6$ thỏa mãn. **1,0 điểm**

IV. Đề số 2b

Bài 1. Điểm kiểm tra môn Toán cuối học kỳ 1 của một nhóm 17 học sinh là:

8, 9, 7, 0, 2, 7, 9, 10, 4, 5, 7, 8, 8, 9, 10, 8, 3.

Tìm một và số trung vị của các số liệu thống kê đã cho.

Lời giải. Sắp xếp số liệu theo thứ tự không giảm ta được:

0, 2, 3, 4, 5, 7, 7, 7, 8, 8, 8, 8, 9, 9, 9, 10, 10.

Suy ra,

- Phần tử có tần số lớn nhất là 8. Do đó một $M_0 = 8$ **1,0 điểm.**
- Trung vị $M_e = 8$ **1,0 điểm**

Bài 2. Tiền thưởng của 35 nhân viên trong một công ti được thống kê trong bảng tần số ghép lớp sau đây (đơn vị: triệu đồng)

Lớp	[20; 24]	[25; 29]	[30; 34]	[35; 39]	[40; 44]	Cộng
Tần số	2	7	15	8	3	$n = 35$

Tính số trung bình và độ lệch chuẩn (kết quả làm tròn đến hàng phần trăm).

Lời giải.

- Số trung bình $\bar{x} = \frac{2 \cdot 22 + 7 \cdot 27 + 15 \cdot 32 + 8 \cdot 37 + 3 \cdot 42}{35} = 32,43$ **1,0 điểm.**

- Phương sai

$$s_x^2 = \frac{2 \cdot 22^2 + 7 \cdot 27^2 + 15 \cdot 32^2 + 8 \cdot 37^2 + 3 \cdot 42^2}{35} - \left(\frac{2 \cdot 22 + 7 \cdot 27 + 15 \cdot 32 + 8 \cdot 37 + 3 \cdot 42}{35} \right)^2.$$

Do vậy độ lệch chuẩn $s_x = \sqrt{s^2} = 4,98$ **1,0 điểm**

Bài 3. Điều tra về cân nặng của cá (đơn vị: kg) trong ao nuôi, người ta thu được bảng số liệu thống kê sau:

3,5	8,5	6,5	5,5	4,5	5,6	9,5	5,9	7,5	3,8
6,8	4,2	7,3	5,7	8,0	5,9	3,6	7,8	6,4	7,9
7,1	6,0	7,6	4,0	6,1	6,7	6,3	6,9	6,2	7,7
6,1	3,8	6,2	6,6	6,3	6,4	7,5	3,9	7,5	6,0
9,0	6,0	7,5	3,6	6,1	6,0	5,4	6,4	9,2	4,7

a) Lập bảng phân bố tần số ghép lớp, sử dụng các lớp sau: $[3,5; 5,0)$, $[5,0; 6,5)$, $[6,5; 8,0)$, $[8,0; 10]$.

b) Vẽ biểu đồ tần số hình cột.

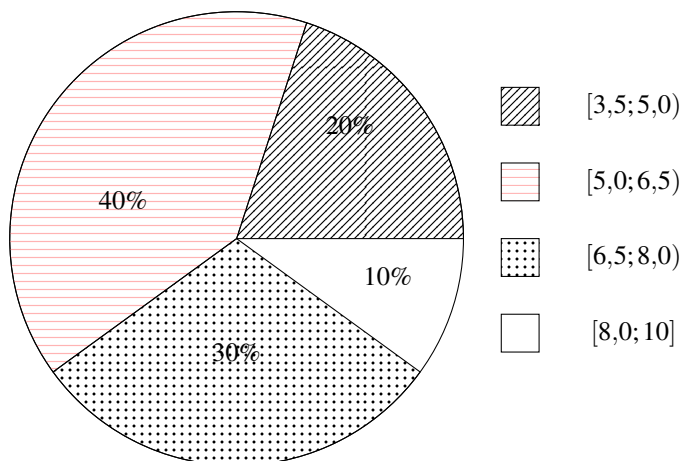
Lời giải.

a) Bảng phân bố tần số ghép lớp:

Lớp	$[3,5; 5,0)$	$[5,0; 6,5)$	$[6,5; 8,0)$	$[8,0; 10]$	Cộng
Tần số	10	20	15	5	$n = 50$

..... **1,0 điểm**

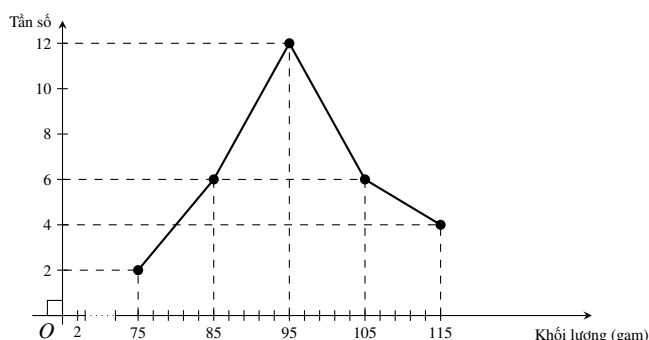
b) Biểu đồ tần số hình quạt:



..... **2,0 điểm**

Bài 4.

Cho biểu đồ đường gấp khúc tần số hình bên mô tả khối lượng của 30 củ khoai tây sau khi thu hoạch (đơn vị: gam). Hãy lập bảng phân bố tần số và tần suất ghép lớp của dữ liệu đó, sử dụng các lớp sau: $[70; 80)$, $[80; 90)$, $[90; 100)$, $[100; 110)$, $[110; 120]$.



Lời giải. Bảng phân bố tần số và tần suất ghép lớp:

Lớp	$[70; 80)$	$[80; 90)$	$[90; 100]$	$[100; 110)$	$[110; 120]$	Cộng
Tần số	2	6	12	6	4	$n = 30$
Tần suất (%)	6,7	20,0	40,0	20,0	13,3	100 (%)

..... **2,0 điểm**

Bài 5. Điểm kiểm tra môn Toán của 40 học sinh lớp 10B được thống kê trong bảng phân bố tần số sau đây (thang điểm 10):

Điểm	0	1	2	3	4	5	6	7	8	9	10	Cộng
Tần số	2	1	2	1	2	x	5	7	y	5	4	$n = 40$

Biết rằng mẫu số liệu trên có 2 một. Hãy tìm x và y .

Lời giải. Tổng số học sinh là 45 nên $x + y = 16$, suy ra có ít nhất một trong hai số x hoặc y không nhỏ hơn 8. Vì mẫu số liệu có 2 một nên $x = y = 8$ thỏa mãn. **1,0 điểm**

V. Đề số 3a

Bài 1. (2,0 điểm) Cho bảng phân bố tần số ghép lớp:

Lớp các giá trị x	$[8; 10)$	$[10; 12)$	$[12; 14]$	Cộng
Tần số n_i	15	30	55	100

Hãy tìm số trung bình của các giá trị trong bảng trên.

Lời giải. Giá trị đại diện của lớp $[8; 10)$: $c_1 = \frac{8+10}{2} = 9$.

Giá trị đại diện của lớp $[10; 12)$: $c_2 = \frac{10+12}{2} = 11$.

Giá trị đại diện của lớp $[12; 14)$: $c_3 = \frac{12+14}{2} = 13$.

Số trung bình cộng: $\bar{x} = \frac{9 \cdot 15 + 11 \cdot 30 + 13 \cdot 55}{15 + 30 + 55} = \frac{59}{5}$.

Bài 2. (2,0 điểm) Độ dài của 60 lá dương xỉ trưởng thành được cho bằng bảng phân bố tần số ghép lớp như sau.

Số TT	Lớp của độ dài (cm)	Tần số
1	$[10; 20)$	18
2	$[20; 30)$	8
3	$[30; 40)$	10
4	$[40; 50)$	24
	Cộng	60

Hãy tính phương sai của bảng số liệu thống kê trên.

Lời giải. Lập thêm cột giá trị đại diện của mỗi lớp:

Số TT	Lớp của độ dài (cm)	Giá trị đại diện	Tần số
1	[10; 20)	15	18
2	[20; 30)	25	8
3	[30; 40)	35	10
4	[40; 50)	45	24
	Cộng		60

Ta có độ dài trung bình là $\bar{x} = \frac{95}{3}$.

Vậy phương sai cần tìm là

$$s_x^2 = \frac{18 \left(15 - \frac{95}{3}\right)^2 + 8 \left(25 - \frac{95}{3}\right)^2 + 10 \left(35 - \frac{95}{3}\right)^2 + 24 \left(45 - \frac{95}{3}\right)^2}{60} = \frac{1460}{9}.$$

Bài 3. (2,0 điểm) Cho bảng phân bố tần số sau

Giá trị	x_1	x_2	x_3	x_4	x_5
Tần số	3	5	$n + 6$	$20 - n$	9

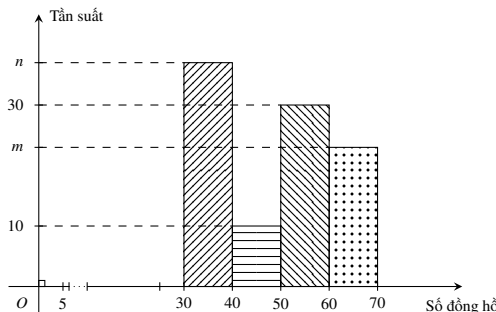
Trong đó n là số tự nhiên và giá trị x_4 là một duy nhất của bảng số liệu thống kê đã cho. Hãy tìm số n ?

Lời giải. Từ giả thiết x_4 là một duy nhất của bảng số liệu thống kê đã cho nên ta có

$$\begin{cases} 20 - n > 9 \\ 20 - n > n + 6 \end{cases} \Leftrightarrow \begin{cases} n < 11 \\ n < 7 \end{cases} \Leftrightarrow n < 7.$$

Vì n là số tự nhiên nên các giá trị n thỏa mãn là: $0 \leq n < 7$.

Bài 4. (2,0 điểm) Cho biểu đồ tần suất hình cột về số đồng hồ bán được của một cửa hàng ở thị trấn X trong 30 ngày như hình vẽ



Tìm giá trị của m và n biết $m - n = 14$.

Lời giải. Dựa vào biểu đồ ta có: $10 + m + 30 + n = 100 \Rightarrow m + n = 60$.

Ta có hệ phương trình: $\begin{cases} m + n = 60 \\ m - n = 14 \end{cases} \Rightarrow \begin{cases} m = 37 \\ n = 23 \end{cases}$.

Bài 5. (2,0 điểm) Một đoàn gồm 80 học sinh của tỉnh A (gồm lớp 11 và lớp 12) tham dự kì thi giải toán của tỉnh (thang điểm 20) và điểm trung bình của họ là 10. Biết rằng số học sinh lớp 11 nhiều hơn 50% số học sinh lớp 12 và điểm trung bình của học sinh khối 12 cao hơn điểm trung bình của học sinh khối 11 là 50%. Hãy tính điểm trung bình của học sinh khối 12.

Lời giải. Gọi x là số học sinh khối 12.

Số học sinh của khối 11 là $1,5x$.

Theo đề: $x + 1,5x = 80 \Leftrightarrow x = 32$.

Suy ra số học sinh của khối 11 là $1,5.32 = 48$ (học sinh)

Gọi y là điểm trung bình của học sinh khối 11.

Suy ra điểm trung bình của học sinh khối 12 là $1,5y$.

Điểm trung bình của 80 học sinh bằng 10 nên ta có: $\frac{48.y + 1,5y.32}{80} = 10 \Leftrightarrow y = \frac{25}{3}$.

Suy ra điểm trung bình của học sinh khối 11 là: $1,5y = 1,5 \cdot \frac{25}{3} = 12,5$.

VI. Đề số 3b

Bài 1. (2,0 điểm) Điểm thi học kì của một học sinh như sau: 5, 5, 2, 6, 3, 5, 9, 8, 7, 10, 9. Tìm số trung bình và số trung vị.

Lời giải. Ta có số trung bình là: $\bar{n} = \frac{5+5+2+6+3+5+9+8+7+10+9}{11} = \frac{69}{11}$.

Xếp lại theo thứ tự không giảm: 2, 3, 5, 5, 5, 6, 7, 8, 9, 9, 10. Có 11 số liệu nên trung vị là số liệu thứ 6. Đó là số 6.

Bài 2. (2,0 điểm) Nhiệt độ trung bình ở tháng 12 của tỉnh X trong suốt 30 năm qua đã được ghi lại theo bảng phân bố tần suất ghép lớp như sau:

Lớp nhiệt độ	Tần suất (%)
[12; 16)	16,70
[16; 20)	43,25
[20; 24)	36,75
[24; 28]	3,30
Cộng	100%

Tìm độ lệch chuẩn của bảng số liệu trên (làm tròn đến chữ số thập phân thứ 2)?

Lời giải. Ta có:

$$\overline{x^2} = \frac{16,70 \cdot 14^2 + 43,25 \cdot 18^2 + 36,75 \cdot 22^2 + 3,30 \cdot 26^2}{100} \approx 373,04.$$

$$\bar{x} = \frac{16,70 \cdot 14 + 43,25 \cdot 18 + 36,75 \cdot 22 + 3,30 \cdot 26}{100} \approx 19,07$$

$$(\bar{x})^2 = 19,07^2 \approx 363,66.$$

$$s_x^2 = 373,04 - 363,66 \approx 9,38 \Rightarrow s_x = \sqrt{9,38} \approx 3,06.$$

Bài 3. (2,0 điểm) Cho bảng phân bố tần số sau

Giá trị	x_1	x_2	x_3	x_4	x_5
Tần số	5	10	n^2	16	$6n - 5$

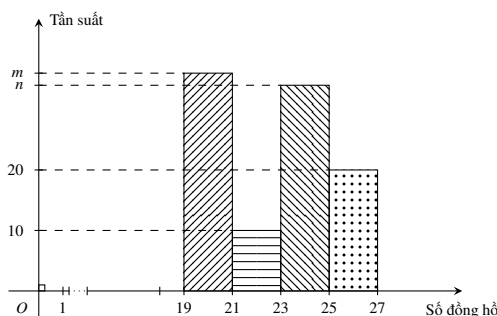
Tìm tất cả các giá trị n để x_3 là một duy nhất của bảng phân bố tần số đã cho.

Lời giải. Từ giả thiết x_3 là một duy nhất của bảng số liệu thống kê đã cho nên ta có

$$\begin{cases} n^2 > 16 \\ n^2 > 6n - 5 \end{cases} \Leftrightarrow \begin{cases} n < -4 \\ n > 4 \\ n < 1 \\ n > 5 \end{cases} \Leftrightarrow \begin{cases} n < -4 \\ n > 5 \end{cases}$$

Vì n là số tự nhiên nên các giá trị n thỏa mãn là: $n > 5$.

Bài 4. (2,0 điểm) Cho biểu đồ tần suất hình cột về số đồng hồ được bán ra của một cửa hàng ở tỉnh X trong 30 ngày như hình vẽ



Tìm giá trị của m và n biết $m^2 - 15m - 250 = 0$.

Lời giải. Ta có: $m^2 - 15m - 250 = 0 \Leftrightarrow m = 25$.

Dựa vào biểu đồ ta có: $10 + 20 + m + n = 100 \Rightarrow m + n = 70 \Rightarrow n = 45$.

Bài 5. (2,0 điểm) Sau một kì thi học sinh giỏi Toán, người ta thống kê kết quả (thang điểm 20) và thu được bảng tần số sau.

Lớp điểm	[6; 10]	[11; 15]	[16; 20]	Cộng
Tần số	22	12	6	40

Nếu những học sinh chỉ cần đạt điểm trung bình của bảng điểm trên đều được nhận Giấy Khen của ban tổ chức, thì số học sinh được nhận Giấy Khen là bao nhiêu?

Lời giải. Ta lập lại bảng với thêm dòng giá trị đại diện:

Lớp điểm	[6; 10]	[11; 15]	[16; 20]	Cộng
Giá trị đại diện	8	13	18	
Tần số	22	12	6	40

Điểm trung bình là: $\bar{x} = \frac{22 \cdot 8 + 12 \cdot 13 + 6 \cdot 18}{40} = 11.$

Số học sinh được nhận thưởng là: $12 + 6 = 18$ (học sinh).