TÀI LIỆU THUỘC KHÓA HỌC
"LIVE VIP 9+ TOÁN "

ĐĂNG KÝ HỌC EM INBOX THẦY TƯ VẤN NHÉ!

TOÁN THỰC TẾ LỚP 12 SGK MỚI
(Chuẩn cấu trúc SGK form mới)

CHUYÊN ĐỀ 6. THỐNG KÊ

A. KIẾN THỨC CƠ BẢN CẦN NẮM

I. KHOẢNG BIẾN THIÊN

Cho mẫu số liệu ghép nhóm:

Nhóm	$[a_1; a_2)$	...	$[a_i; a_{i+1})$	...	$[a_k; a_{k+1})$
Tần số	m_1	...	m_i	...	m_k

trong đó các tần số $m_1 > 0, m_k > 0$ **và** $n = m_1 + \ldots + m_k$ **là cỡ mẫu.**

Khoảng biến thiên của mẫu ghép nhóm trên là $R = a_{k+1} - a_1$.

Ý nghĩa. Khoảng biến thiên của mẫu số liệu ghép nhóm xấp xỉ cho khoảng biến thiên của mẫu số liệu gốc. Khoảng biến thiên được dùng để đo mức độ phân tán của mẫu số liệu ghép nhóm. Khoảng biến thiên càng lớn thì mẫu số liệu càng phân tán.

Lưu ý: Khoảng biến thiên của MSL ghép nhóm luôn ≥ khoảng biến thiên của MSL gốc.

II. KHOẢNG TỨ PHÂN VỊ

Tứ phân vị thứ r là

$$Q_r = a_p + \frac{\frac{r \cdot n}{4} - \left(m_1 + \ldots + m_{p-1}\right)}{m_p} \cdot \left(a_{p+1} - a_p\right),$$

trong đó $\left[a_p; a_{p+1}\right)$ là nhóm chứa tứ phân vị thứ r với $r = 1, 2, 3$.

Khoảng tứ phân vị của mẫu số liệu ghép nhóm, kí hiệu là Δ_Q, là hiệu số giữa tứ phân vị thứ ba Q_3 và tứ phân vị thứ nhất Q_1 của mẫu số liệu đó, tức là $\Delta_Q = Q_3 - Q_1$.

Ý nghĩa. Khoảng tứ phân vị của mẫu số liệu ghép nhóm xấp xỉ cho khoảng tứ phân vị của mẫu số liệu gốc. Khoảng tứ phân vị cũng được dùng để đo mức độ phân tán của mẫu số liệu ghép nhóm. Khoảng tứ phân vị càng lớn thì mẫu số liệu càng phân tán.

Nhận xét.

- Do khoảng tứ phân vị của mẫu số liệu ghép nhóm chỉ phụ thuộc vào nửa giữa của mẫu số liệu, nên không bị ảnh hưởng bởi các giá trị bất thường và có thể dùng đại lượng này để loại giá trị bất thường.
- Khoảng tứ phân vị được dùng để xác định giá trị ngoại lệ trong mẫu số liệu. Giá trị x trong mẫu số liệu là giá trị ngoại lệ nếu $x > Q_3 + 1,5\Delta_Q$ hoặc $x < Q_1 - 1,5\Delta_Q$

III. PHƯƠNG SAI VÀ ĐỘ LỆCH CHUẨN

1. Phương sai

Phương sai của mẫu số liệu ghép nhóm, kí hiệu S^2, được tính bởi công thức:

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

Trong đó: $n=n_1+n_2+\ldots+n_k$ là cỡ mẫu;

$\overline{x}=\frac{1}{n}\left(n_1c_1+n_2c_2+\ldots+n_kc_k\right)$ là số trung bình

2. Độ lệch chuẩn

Độ lệch chuẩn của mẫu số liệu ghép nhóm kí hiệu S,

Độ lệch chuẩn là căn bậc hai số học của phương sai:

Độ lệch chuẩn $=\sqrt{S^2}$

Chú ý:

- Phương sai của mẫu số liệu ghép nhóm có thể được tính theo công thức sau:

$$S^2=\frac{1}{n}\left(n_1c_1^2+n_2c_2^2+\ldots+n_kc_k^2\right)-\overline{x}^2$$

- Trong thống kê, người ta còn dùng đại lượng sau để đo mức độ phân tán của mẫu số liệu ghép nhóm:

$$S^2=\frac{1}{n-1}\left[n_1\left(c_1-\overline{x}\right)^2+n_2\left(c_2-\overline{x}\right)^2+\ldots+n_k\left(c_k-\overline{x}\right)^2\right].$$

Ý nghĩa:

- ***Phương sai*** của MSL ghép nhóm là giá trị xấp xỉ cho phương sai của MSL gốc.

Độ lệch chuẩn của MSL ghép nhóm cũng là giá trị xấp xỉ cho độ lệch chuẩn của MSL gốc.

$\longrightarrow$ Được dùng để đo mức độ phân tán của mẫu số liệu ghép nhóm.

- Phương sai và độ lệch chuẩn càng lớn thì dữ liệu càng phân tán.
- Độ lệch chuẩn có cùng đơn vị với đơn vị của mẫu số liệu.

B. BÀI TẬP VẬN DỤNG

Câu 1: Dữ liệu về tốc độ của 100 xe ô tô lưu thông trên một đoạn đường cao tốc vào giờ cao điểm, được trích xuất từ camera của cơ quan cảnh sát giao thông. Hãy tìm khoảng biến thiên của mẫu số liệu (bảng số liệu hình bên dưới).

Tốc độ (km/h)	$[60;70)$	$[70;80)$	$[80;90)$	$[90;100)$	$[100;110)$
Số xe	10	20	20	35	15

Câu 2: Thời gian hoàn thành bài kiểm tra của các bạn trong lớp 12A được cho bảng sau:

Thời gian (phút)	$[25;30)$	$[30;35)$	$[35;40)$	$[40;45)$
Số học sinh	8	16	4	2

a) Tính khoảng biến thiên R cho mẫu số liệu ghép nhóm trên.

b) Nếu biết học sinh hoàn thành bài kiểm tra sớm nhất mất 27 phút và muộn nhất mất 43 phút. Hãy so sánh khoảng biến thiên của mẫu số liệu ghép nhóm và mẫu số liệu gốc.

Câu 3: Để chuẩn bị mở một trung tâm thể dục thể thao, anh Tiến đã tiến hành điều tra tuổi thọ của máy chạy bộ (đơn vị: năm) do hai hãng X, Y sản xuất. Bảng biểu thị hai mẫu số liệu mà anh thu thập được qua Internet.

Tuổi thọ	$[2;4)$	$[4;6)$	$[6;8)$	$[8;10)$	$[10;12)$
Số máy của hãng X	7	20	36	20	17
Số máy của hãng Y	0	20	35	35	10

Khoảng biến thiên của mẫu số liệu nào lớn hơn? Từ đó có thể nói là máy chạy bộ do hãng nào sản xuất có tuổi thọ phân tán hơn?

Câu 4: Bảng dưới biểu thị kết quả điều tra thời gian sử dụng Internet hằng ngày của một số người.

Thời gian (phút)	$[30;60)$	$[60;90)$	$[90;120)$	$[120;150)$	$[150;180)$
Số người	2	4	10	5	3

Tìm khoảng biến thiên của mẫu số liệu đã cho. Kết quả cho biết điều gì?

Câu 5: Người ta tiến hành phỏng vấn hai nhóm khán giả về một bộ phim mới công chiếu. Nhóm A gồm những khán giả thuộc lứa tuổi 20 - 30, nhóm B thuộc lứa tuổi trên 30. Người được hỏi ý kiến phải đánh giá bộ phim bằng cách cho điểm theo một số tiêu chí nêu trong phiếu điều tra và sau đó lấy tổng số điểm (thang điểm 100). Bảng dưới đây trình bày kết quả điều tra hai nhóm khán giả:

Bảng. Điểm đánh giá của khán giả

Điểm	$[50;60)$	$[60;70)$	$[70;80)$	$[80;90)$	$[90;100)$
Số người của nhóm A	6	10	14	12	8
Số người của nhóm B	0	8	14	28	0

Ý kiến đánh giá của nhóm khán giả nào phân tán hơn?

Câu 6: Một người ghi lại thời gian đàm thoại của một số cuộc gọi cho kết quả như bảng sau:

Thời gian t (phút)	$0 \le t < 1$	$1 \le t < 2$	$2 \le t < 3$	$3 \le t < 4$	$4 \le t < 5$
Số cuộc gọi	8	17	25	20	10

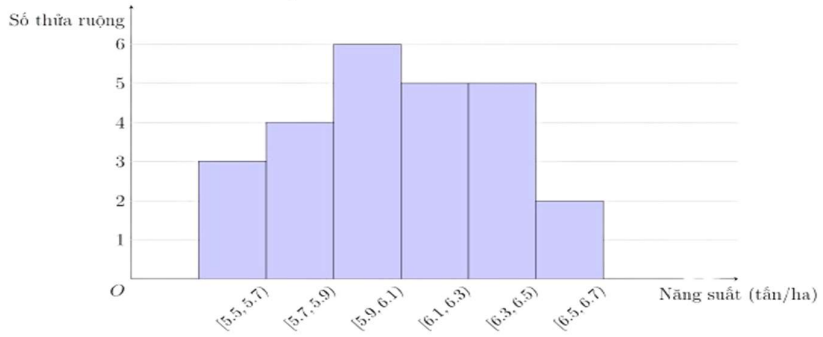
Tính khoảng tứ phân vị của mẫu số liệu ghép nhóm trên.

Câu 7: Thầy Cư rất thích chạy bộ. Thời gian chạy bộ mỗi ngày trong thời gian gần đây của Thầy Cư được thống kê lại ở bảng sau:

Thời gian (phút)	$[20;25)$	$[25;30)$	$[30;35)$	$[35;40)$	$[40;45)$
Số ngày	6	6	4	1	1

Hãy tính khoảng tứ phân vị của mẫu số liệu ghép nhóm trong bảng trên.

Câu 8: Khảo sát năng suất của một số thửa ruộng được minh họa ở biểu đồ sau:



a) Có bao nhiêu thửa ruộng được khảo sát?

b) Lập bảng tần số ghép nhóm

c) Hãy xác định khoảng biến thiên và khoảng tứ phân vị của mẫu số liệu trên.

Câu 9: cho kết quả như sau:

101 79 79 78 75 73 68 67 67 63
63 61 60 59 57 55 55 50 47 42.

(Theo *premierleague.com*)

a) Hãy ghép nhóm dãy số liệu trên thành các nhóm có độ dài bằng nhau với nhóm đầu tiên là $[40;50)$.

b) Tính khoảng biến thiên, khoảng tứ phân vị của mẫu số liệu gốc và mẫu số liệu ghép nhóm thu được ở câu a. Giá trị nào là giá trị chính xác? Giá trị nào là giá trị xấp xỉ?

Câu 10: Thu nhập theo tháng (đơn vị: triệu đồng) của người lao động ở hai nhà máy như sau:

Thu nhập	[5;8)	[8;11)	[11;14)	[14;17)	[17;20)
Số người của nhà máy A	20	35	45	35	20
Số người của nhà máy B	17	23	30	23	17

Tính mức thu nhập trung bình của người lao động ở hai nhà máy trên. Dựa vào khoảng tứ phân vị, hãy xác định xem mức thu nhập của người lao động ở nhà máy nào biến động nhiều hơn.

Câu 11: Bảng sau đây cho biết chiều cao của các học sinh lớp $12A$ và $12B$.

Chiều cao (cm)	[145; 150)	[150; 155)	[155; 160)	[160; 165)	[165; 170)	[170; 175)
Số học sinh của lớp 12A	1	0	15	12	10	5
Số học sinh của lớp 12B	0	0	17	10	9	6

a) Tìm khoảng biến thiên, khoảng tứ phân vị cho các mẫu số liệu ghép nhóm về chiều cao của học sinh lớp 12A, 12B

b) Để so sánh độ phân tán về chiều cao của học sinh hai lớp này ta nên dùng khoảng biến thiên hay khoảng tứ phân vị? Vì sao?

Câu 12: Điểm kiểm tra cuối khoá môn Tiếng Anh của hai lớp ở một trung tâm ngoại ngữ được thống kê trong các Bảng a và b

Bảng a. Điểm của lớp A

Điểm	Số học viên (tần số)
[50;60)	8
[60;70)	20
[70;80)	50
[80;90)	17
[90;100)	5

Bảng b. Điểm của lớp B

Điểm	Số học viên (tần số)
[50;60)	15
[60;70)	20
[70;80)	30
[80;90)	20
[90;100)	15

a) Tìm khoảng biến thiên của mỗi mẫu số liệu. Có thể dùng khoảng biến thiên để biết điểm của lớp nào đồng đều hơn không?

b) Tìm các tứ phân vị và khoảng tứ phân vị của mỗi mẫu số liệu.

c) Mẫu số liệu nào có độ phân tán lớn hơn? Minh hoạ câu trả lời bằng cách biểu diễn các tứ phân vị và khoảng tứ phân vị của mỗi mẫu số liệu trên trục số.

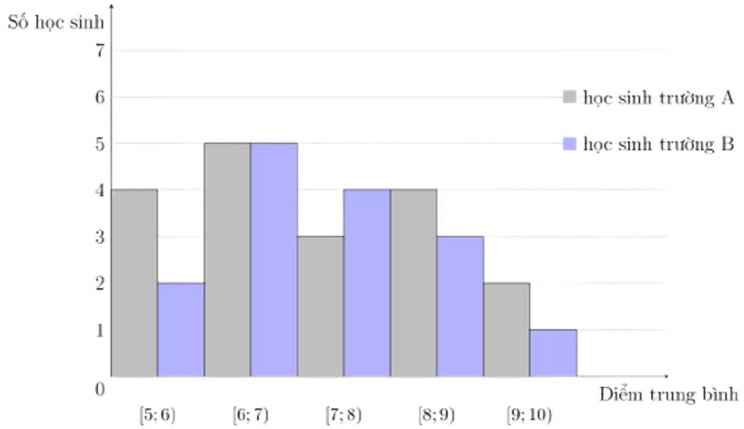
Câu 13: Kết quả đo chiều cao của 100 cây dừa trồng sau 10 năm tại một vườn trái cây ở Bến Tre cho ở bảng sau:

Chiều cao (m)	$[8,4;8,6)$	$[8,6;8,8)$	$[8,8;9,0)$	$[9,0;9,2)$	$[9,2;9,4)$
Số cây	5	12	25	44	14

a) Hãy tìm khoảng biến thiên, khoảng tứ phân vị của mẫu số liệu ghép nhóm trên.

b) Trong 100 cây dừa trên có 1 cây cao 8,4 m. Hỏi chiều cao của cây dừa này có phải là giá trị ngoại lệ không?

Câu 14: Biểu đồ sau mô tả kết quả điều tra về điểm trung bình năm học của học sinh hai trường A và *B*.



a) Hãy xác định giá trị đại diện cho mỗi nhóm và lập bảng tần số ghép nhóm cho mẫu số liệu trên.

b) Nếu so sánh theo khoảng tứ phân vị của mẫu số liệu ghép nhóm thì học sinh trường nào có điểm trung bình đều hơn?

Câu 15: Thầy giáo thống kê lại điểm trung bình cuối năm của các học sinh lớp 11 A và 11B ở bảng sau:

Điểm trung bình	$[5;6)$	$[6;7)$	$[7;8)$	$[8;9)$	$[9;10)$
11A	1	0	11	22	6
11B	0	6	8	14	12

a) Nếu so sánh theo khoảng biến thiên thì học sinh lớp nào có điểm trung bình ít phân tán hơn?

b) Nếu so sánh theo độ lệch chuẩn thì học sinh lớp nào có điểm trung bình ít phân tán hơn?

Câu 16: Cân nặng (kg) của một số quả mít trong một khu vườn được thống kê ở bảng sau:

Cân nặng (kg)	$[4;6)$	$[6;8)$	$[8;10)$	$[10;12)$	$[12;14)$
Số cây giống	6	12	19	9	4

Hãy tính phương sai của mẫu số liệu ghép nhóm trên.

Câu 17: Thống kê điểm trung bình của hai lớp 12A và 12B được cho ở bảng sau

Điểm trung bình	$[5;6)$	$[6;7)$	$[7;8)$	$[8;9)$	$[9;10)$
Số học sinh lớp 12A	1	0	11	22	6
Số học sinh lớp 12B	0	6	8	14	12

Nếu so sánh phương sai thì học sinh lớp nào có điểm trung bình ít phân tán hơn?

Câu 18: Tuổi thọ của một số linh kiện điện tử (đơn vị: năm) được sản xuất bởi hai phân xưởng được cho như sau:

Tuổi thọ (năm)	$[1,5;2)$	$[2;2,5)$	$[2,5;3)$	$[3;3,5)$	$[3,5;4)$
Số linh kiện của phân xưởng 1	4	9	13	8	6
Số linh kiện của phân xưởng 2	2	8	20	7	3

Tính phương sai của mỗi mẫu số liệu ghép nhóm và nhận xét về độ phân tán của tuổi thọ các linh kiện điện tử được sản xuất bởi mỗi phân xưởng.

Câu 19: Cho mẫu số liệu ghép nhóm về độ tuổi của cư dân trong một khu phố.

Độ tuổi	$[20;30)$	$[30;40)$	$[40;50)$	$[50;60)$	$[60;70)$	$[70;80)$
Giá trị đại diện	25	35	45	55	65	75
Số cư dân	25	20	20	15	14	6

Tính phương sai và độ lệch chuẩn của mẫu số liệu đó. (*Kết quả làm tròn đến hàng phần mười*)

Câu 20: Cho hai mẫu số liệu ghép nhóm thống kê mức lương của hai công ty A, B (đơn vị là triệu đồng)

Bảng thống kê mức lương công ty A (triệu đồng)	
Mức lương	Số lượng nhân viên
$[10;15)$	15
$[15;20)$	18
$[20;25)$	10
$[25;30)$	10
$[30;35)$	5
$[35;40)$	2

Bảng thống kê mức lương công ty B (triệu đồng)	
Mức lương	Số lượng nhân viên
$[10;15)$	25
$[15;20)$	15
$[20;25)$	7
$[25;30)$	5
$[30;35)$	5
$[35;40)$	3

a) Tính phương sai và độ lệch chuẩn của mẫu số liệu ghép nhóm lần lượt biểu diễn mức lương của hai công ty A, B. (*Kết quả làm tròn đến hàng phần mười*)

b) Công ty nào có mức lương đồng đều hơn?

Câu 21: Giá đóng cửa của một cổ phiếu là giá của cổ phiếu đó cuối một phiên giao dịch. Bảng sau thống kê giá đóng của (đơn vị là nghìn đồng) của hai mã cổ phiếu A và B trong 50 ngày giao dịch liên tiếp.

Giá đóng cửa	$[120;122)$	$[122;124)$	$[124;126)$	$[126;128)$	$[128;130)$
Cổ phiếu A	9	8	11	13	9
Cổ phiếu B	14	6	5	20	5

a) Tính phương sai và độ lệch chuẩn của hai mã cổ phiếu A, B. (*Kết quả làm tròn đến hàng phần mười*)

b) So sánh độ rủi ro của cổ phiếu A và B?

Câu 22: Chỉ số P/B (viết tắt của từ Price to book ratio) là một chỉ số tài chính quan trọng, được sử dụng để so sánh giữa giá của cổ phiếu với giá trị ghi sổ của cổ phiếu đó. Khi chỉ số P/B cao, nghĩa là thị trường đang có nhiều kỳ vọng đối với cổ phiếu này, doanh nghiệp có tiềm năng phát triển tốt trong tương lai. Bảng sau thống kê chỉ số P/B của hai mã cổ phiếu A và B trong 10 năm giao dịch liên tiếp.

Giá đóng cửa	$[0;2)$	$[2;4)$	$[4;6)$	$[6;8)$	$[8;10)$	$[10;12)$
Cổ phiếu A	1	0	4	2	2	1
Cổ phiếu B	1	2	1	2	3	1

a) Tính phương sai và độ lệch chuẩn của hai mã cổ phiếu A, B. (*Kết quả làm tròn đến hàng phần mười*)

b) So sánh độ rủi ro của cổ phiếu A và B? Biết rằng các chỉ số tài chính còn lại của hai cổ phiếu đó là xấp xỉ nhau.

Câu 23: Thời gian hoàn thành một bài viết chính tả của một số học sinh lớp 4 hai trường X và Y được ghi lại ở bảng sau:

Thời gian (phút)	[6;7)	[7;8)	[8;9)	[9;10)	[10;11)
Học sinh trường X	8	10	13	10	9
Học sinh trường Y	4	12	17	14	3

a) Nếu so sánh theo số trung bình thì học sinh trường nào viết nhanh hơn?

b) Nếu so sánh theo khoảng tứ phân vị thì học sinh trường nào có tốc độ viết đồng đều hơn?

c) Nếu so sánh theo độ lệch chuẩn thì học sinh trường nào có tốc độ viết đồng đều hơn?

Câu 24: Một nhóm 20 học sinh dùng một thiết bị đo đường kính của một nhân tế bào cho kết quả như sau:

Kết quả đo (μm)	[4,5;5)	[5;5,5)	[5,5;6)	[6;6,5)
Số học sinh	3	8	7	2

a) Tính số trung bình và độ lệch chuẩn của mẫu số liệu ghép nhóm trên.
b) Số trung bình và độ lệch chuẩn cho biết thông tin gì?

Câu 25: Thời gian chạy tập luyện cự li 100 m của hai vận động viên được cho trong bảng sau:

Thời gian (giây)	[10;10,3)	[10,3;10,6)	[10,6;10,9)	[10,9;11,2)
Số lần chạy của A	2	10	5	3
Số lần chạy của B	3	7	9	6

Dựa trên độ lệch chuẩn của các mẫu số liệu ghép nhóm, hãy cho biết vận động viên nào có thành tích luyện tập ổn định hơn.

Câu 26: Mai và Ngọc cùng sử dụng vòng đeo tay thông minh để ghi lại số bước chân hai bạn đi mỗi ngày trong một tháng. Kết quả được ghi lại ở bảng sau:

Số bước (đơn vị: nghìn)	[3;5)	[5;7)	[7;9)	[9;11)	[11;13)
Mai	6	7	6	6	5
Ngọc	2	5	13	8	2

a) Hãy tính số trung bình và độ lệch chuẩn của mẫu số liệu ghép nhóm trên.
b) Nếu so sánh theo độ lệch chuẩn thì bạn nào có số lượng bước chân đi mỗi ngày đều đặn hơn?

Câu 27: Tốc độ của 20 xe hơi khi đi qua một trạm kiểm tra tốc độ (đơn vị: km/h) được thống kê lại như sau:

42	43,4	43,4	46,5	46,7	46,8	47,5	47,7	48,1	48,4
50,8	52,1	52,7	53,9	54,8	55,6	57,5	59,6	60,3	61,1

a) Hãy tính khoảng biến thiên, khoảng tứ phân vị và độ lệch chuẩn của mẫu số liệu trên.
b) Hãy lập bảng tần số ghép nhóm với nhóm đầu tiên là [42;46) và độ dài mỗi nhóm bằng 4.
c) Hãy tính khoảng biến thiên, khoảng tứ phân vị và độ lệch chuẩn của mẫu số liệu ghép nhóm.

Câu 28: Điều tra thời gian phải làm thêm trung bình hằng tuần của các bác sĩ ở một bệnh viện, người ta thu được số liệu sau:

5	6	7	8	8	8	9	9	9
9	9	10	10	10	10	11	11	11
12	12	12	13	13	14	14		

a) Chuyển mẫu số liệu đã cho về mẫu số liệu ghép nhóm với độ dài các nhóm ghép bằng 2 và nhóm đầu tiên là [5;7).
b) Tính số trung bình, phương sai, độ lệch chuẩn của mẫu số liệu ghép nhóm lập ở câu a (làm tròn kết quả đến hàng phần trăm).
c) So sánh và nêu ý nghĩa các kết quả tìm được ở câu b với các kết quả tương ứng của mẫu số liệu gốc.

Quý phụ huynh và các em học sinh đky khóa học

1. **Zalo thầy Thuận Toán: 0869998668**
   zalo.me/84869998668
2. **Facebook thầy:**
   https://www.facebook.com/Thaygiaothuan.99/
3. **Fanpage thầy:**
   https://www.facebook.com/thaythuantoan
4. **Tiktok thầy:**
   https://www.tiktok.com/@thay_hothucthuan
5. **Kênh Youtube thầy:**
   https://www.youtube.com/@HoThucThuanOfficial