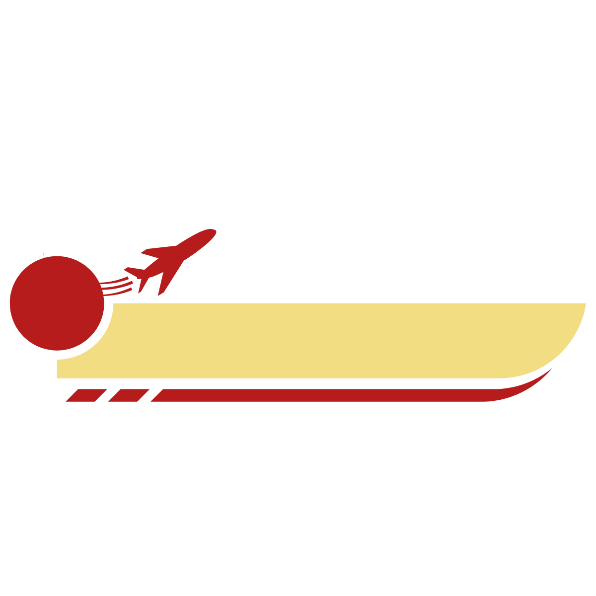
**KHOẢNG BIẾN THIÊN – KHOẢNG TỨ PHÂN VỊ CỦA MSLGN**



**Bài 1.**

**Chương 03**



**A**

**Lý thuyết**

# 1. Khoảng biến thiên.

**Định nghĩa:**

Khoảng biến thiên, kí hiệu , của mẫu số liệu ghép nhóm là hiệu số giữa đầu mút phải của nhóm cuối cùng và đầu mút trái của nhóm đầu tiên có chứa dữ liệu của mẫu số liệu.



» Xét mẫu số liệu ghép nhóm được cho ở bảng sau:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Nhóm |  |  |  |  |
| Tần số |  |  |  |  |

» Nếu  và  cùng khác 0 thì .

▪Khoảng biến thiên của mẫu số liệu ghép nhóm luôn *lớn hơn* hoặc *bằng* khoảng biến thiên của mẫu số liệu gốc.

▪Khoảng biến thiên của mẫu số liệu ghép nhóm là giá trị xấp xỉ khoảng biến thiên của mẫu số liệu gốc và có thể dùng để đo mức độ phân tán của mẫu số liệu.

▪Khoảng biến thiên  chưa phản ánh được đầy đủ mức độ phân tán của phần lớn các số liệu.

▪Hơn nữa, giá trị của  thường tăng vọt khi xuất hiện giá trị ngoại lệ trong mẫu số liệu. Do đó, để phản ánh mức độ phân tán của số liệu, người ta còn dùng các số đặc trưng khác.

**Ý nghĩa**



# 2. Khoảng tứ phân vị

**Định nghĩa:**

Khoảng tứ phân vị của mẫu số liệu ghép nhóm cho ở Bảng 1, kí hiệu , là hiệu giữa tứ phân vị thứ ba  và tứ phân vị thứ nhất  của mẫu số liệu ghép nhóm đó, tức là:





▪ Khoảng tứ phân vị của mẫu số liệu ghép nhóm là giá trị xấp xỉ cho khoảng tứ phân vị của mẫu số liệu gốc và có thể dùng để đo mức độ phân tán của nửa giữa của mẫu số liệu (tập hợp gồm  số liệu nằm chính giữa mẫu số liệu).

▪ Khoảng tứ phân vị của mẫu số liệu ghép nhóm càng nhỏ thì dữ liệu càng tập trung xung quanh trung vị.

▪ Khoảng tứ phân vị được dùng để xác định giá trị ngoại lệ trong mẫu số liệu.

Giá trị  trong mẫu số liệu là giá trị ngoại lệ nếu  hoặc 

▪ Khoảng tứ phân vị của mẫu số liệu ghép nhóm không bị ảnh hưởng nhiều bởi các giá trị ngoại lệ trong mẫu số liệu.

**Ý nghĩa**



▪ Tứ phân vị thứ , kí hiệu là  với  của mẫu số liệu ghép nhóm được xác định như sau: .

Trong đó:

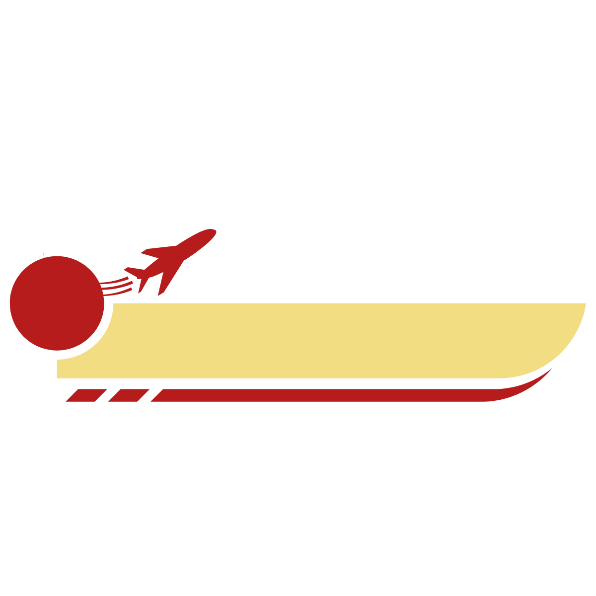
»  là cỡ mẫu.

»  là nhóm chứa tứ phân vị thứ .

»  là tần số của nhóm chứa tứ phân vị thứ .

» 

**Ý nghĩa**



**B**

**Các dạng bài tập**

## 🏱 Dạng 1. Khoảng biến thiên của MSL ghép nhóm

» Xét mẫu số liệu ghép nhóm được cho ở bảng sau:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Nhóm |  |  |  |  |
| Tần số |  |  |  |  |

» Nếu  và  cùng khác 0 thì .

» Khoảng biến thiên của MSL ghép nhóm luôn  khoảng biến thiên của MSL gốc.

***Phương pháp***



**Ví dụ 1.1.**

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) |  |  |  |  |  |
| Số xe |  |  |  |  |  |



🖎 ***Lời giải***

Ta có  và 

Khoảng biến thiên .

**Ví dụ 1.2.**

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

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Thời gian (phút) |  |  |  |  |
| Số học sinh |  |  |  |  |

**⑴** Tính khoảng biến thiên  cho mẫu số liệu ghép nhóm trên.

**⑵** 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.



🖎 ***Lời giải***

**⑴** Tính khoảng biến thiên  cho mẫu số liệu ghép nhóm trên.

Khoảng biến thiên  phút.

**⑵** Hãy so sánh khoảng biến thiên của mẫu số liệu ghép nhóm và mẫu số liệu gốc.

Khoảng biến thiên của mẫu số liệu gốc là  phút.

Vậy khoảng biến thiên của mẫu số liệu ghép nhóm nhỏ hơn khoảng biến thiên của mẫu số liệu gốc.

## 🏱 Dạng 2. Ý nghĩa của khoảng biến thiên trong việc đo mức độ phân tán

» Khoảng biến thiên của MSL ghép nhóm luôn  khoảng biến thiên của MSL gốc.

» Khoảng biến thiên của mẫu số liệu ghép nhóm là giá trị xấp xỉ khoảng biến thiên của mẫu số liệu gốc và có thể dùng để đo mức độ phân tán của mẫu số liệu.

» Khoảng biến thiên  chưa phản ánh được đầy đủ mức độ phân tán của phần lớn các số liệu.

» Hơn nữa, giá trị của  thường tăng vọt khi xuất hiện giá trị ngoại lệ trong mẫu số liệu. Do đó, để phản ánh mức độ phân tán của số liệu, người ta còn dùng các số đặc trưng khác.

***Phương pháp***



**Ví dụ 2.1.**

Để chuẩn bị mở một trung tâm thể dục thể thao, anh Sơn 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 |  |  |  |  |  |
| 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?



🖎 ***Lời giải***

Khoảng biến thiên của tuổi thọ máy chạy bộ do hãng X: 

Khoảng biến thiên của tuổi thọ máy chạy bộ do hãng Y: 

Vì  nên có thể nói là máy do hãng X sản xuất có tuổi thọ phân tán hơn so với máy của hãng Y.

**Ví dụ 2.2.**

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 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:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Tuổi thọ |  |  |  |  |  |
| Số người của nhóm A | 7 | 20 | 36 | 20 | 17 |
| Số người của nhóm B | 0 | 20 | 35 | 35 | 10 |

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



🖎 ***Lời giải***

Khoảng biến thiên của điểm trong phiếu điều tra của số người nhóm A: 

Khoảng biến thiên của điểm trong phiếu điều tra của số người nhóm B: 

Vì  nên có thể nói là số điểm đánh giá theo một tiêu chí của nhóm A phân tán hơn so với nhóm B.

**Ví dụ 2.3.**

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

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Thời gian (phút) |  |  |  |  |  |
| Số người | 2 | 4 | 10 | 5 | 3 |

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



🖎 ***Lời giải***

Đầu mút phải của nhóm ghép cuối cùng là 180,

Đầu mút trái của nhóm ghép đầu tiên là 30.

Vậy khoảng biến thiên của mẫu số liệu là .

Kết quả này cho biết thời gian sử dụng Internet hằng ngày của các thành viên thuộc nhóm người được điều tra chênh lệch nhau nhiều nhất là 150 phút.

## 🏱 Dạng 3. Ý nghĩa của khoảng biến thiên trong việc đo mức độ phân tán

Tứ phân vị thứ *k*, kí hiệu là  với  của MSL ghép nhóm được xác định như sau:



Trong đó:

»  là cỡ mẫu.

»  là nhóm chứa tứ phân vị thứ .

»  là tần số của nhóm chứa tứ phân vị thứ .

» 

» Khoảng tứ phân vị của MSL ghép nhóm cũng được xác định dựa trên tứ phân vị thứ nhất và tứ phân vị thứ ba như đối với MSL không ghép nhóm.

» Khoảng tứ phân vị của MSL ghép nhóm có kí hiệu , là hiệu giữa tứ phân vị thứ ba  và tứ phân vị thứ nhất  của MSL ghép nhóm đó, tức là .

**▪Bước 1:** Xác định cỡ mẫu, tính tứ phân vị thứ nhất và tứ phân vị thứ ba của mẫu số liệu.

**▪Bước 2:** Tính khoảng tứ phân vị của mẫu số liệu.

***Phương pháp***



**Ví dụ 3.1.**

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

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Thời gian  (phút) |  |  |  |  |  |
| Số cuộc gọi | 8 | 17 | 25 | 20 | 10 |

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



🖎 ***Lời giải***

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Thời gian  (phút) |  |  |  |  |  |
| Số cuộc gọi | 8 | 17 | 25 | 20 | 10 |

Cỡ mẫu . Giả sử  là thời gian đàm thoại của 80 cuộc gọi và giả sử dãy số liệu gốc này đã được sắp xếp theo thứ tự tăng dần.

Vì  và  nên nhóm chứa tứ phân vị thứ nhất là nhóm  và tứ phân vị thứ nhất là: 

vi  và  nên nhóm chứa tứ phân vị thứ ba là nhóm  và tứ phân vị thứ ba là: 

Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm là: 

**Ví dụ 3.2.**

Bạn An rất thích chạy bộ. Thời gian chạy bộ mỗi ngày trong thời gian gần đây của bạn An được thống kê lại ở bảng sau:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Thời gian (phút) |  |  |  |  |  |
| Số ngày | 6 | 6 | 4 | 1 | 1 |

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



🖎 ***Lời giải***

Cỡ mẫu .

Gọi  là mẫu số liệu gốc gồm thời gian của 18 ngày chạy bộ của bạn An được sắp xếp theo thứ tự không giảm.

Ta có: 

Tứ phân vị thứ nhất của mẫu số liệu gốc là .

Do đó, tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là:

.

Tứ phân vị thứ ba của mẫu số liệu gốc là .

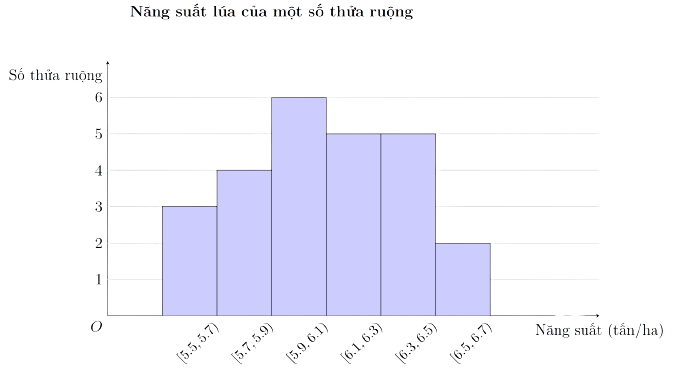
Do đó, tứ phân vị thứ ba của mẫu số liệu ghép nhóm là:

.

Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm là: .

**Ví dụ 3.2.**

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



**⑴** Có bao nhiêu thửa ruộng được khảo sát?

**⑵** Lập bảng tần số ghép nhóm và tần số tương đối ghép nhóm tương ứng của mẫu số liệu trên.

**⑶** Hãy xác định khoảng biến thiên và khoảng tứ phân vị của mẫu số liệu trên.



🖎 ***Lời giải***

**⑴** Có bao nhiêu thửa ruộng được khảo sát?

Có tất cả  thửa ruộng được khảo sát.

**⑵** Lập bảng tần số ghép nhóm và tần số tương đối ghép nhóm tương ứng của mẫu số liệu trên.

Bảng tần số ghép nhóm và tần số tương đối ghép nhóm tương ứng của mẫu số liệu trên.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Năng suất (tấn/ha) |  |  |  |  |  |  |
| Số thửa ruộng | 3 | 4 | 6 | 5 | 5 | 2 |
| Tần số tương đối |  |  |  |  |  |  |

**⑶** Hãy xác định khoảng biến thiên và khoảng tứ phân vị của mẫu số liệu trên.

+ Khoảng biến thiên của mẫu số liệu trên là: .

+ Cỡ mẫu .

Gọi  là mẫu số liệu gốc gồm năng suất của 25 thửa ruộng được sắp xếp theo thứ tự không giảm.

Ta có: ; .

Tứ phân vị thứ nhất của mẫu số liệu gốc là .

Do đó, tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là:

.

Tứ phân vị thứ ba của mẫu số liệu gốc là .

Do đó, tứ phân vị thứ ba của mẫu số liệu ghép nhóm là:

.

Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm là: .

## 🏱 Dạng 4. Ý nghĩa của khoảng tứ phân vị trong việc đo mức độ phân tán

▪ Khoảng tứ phân vị của mẫu số liệu ghép nhóm là giá trị xấp xỉ cho khoảng tứ phân vị của mẫu số liệu gốc và có thể dùng để đo mức độ phân tán của nửa giữa của mẫu số liệu (tập hợp gồm  số liệu nằm chính giữa mẫu số liệu).

▪ Khoảng tứ phân vị của mẫu số liệu ghép nhóm càng nhỏ thì dữ liệu càng tập trung xung quanh trung vị.

▪ Khoảng tứ phân vị được dùng để xác định giá trị ngoại lệ trong mẫu số liệu.

Giá trị  trong mẫu số liệu là giá trị ngoại lệ nếu  hoặc 

▪ Khoảng tứ phân vị của mẫu số liệu ghép nhóm không bị ảnh hưởng nhiều bởi các giá trị ngoại lệ trong mẫu số liệu.

***Phương pháp***



**Ví dụ 4.1.**

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) |  |  |  |  |  |
| Số cây | 5 | 12 | 25 | 44 | 14 |

**⑴** Hãy tìm khoảng biến thiên, khoảng tứ phân vị của mẫu số liệu ghép nhóm trên.

**⑵** 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?



🖎 ***Lời giải***

**⑴** Hãy tìm khoảng biến thiên, khoảng tứ phân vị của mẫu số liệu ghép nhóm trên.

Khoảng biến thiên của mẫu số liệu là .

Ta có cỡ mẫu là .

Gọi  là mẫu số liệu gồm chiều cao của 100 cây dừa.

Ta có:



Tứ phân vị thứ nhất của mẫu số liệu là .

Do đó, tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là

.

Tứ phân vị thứ ba của mẫu số liệu là .

Do đó, tứ phân vị thứ ba của mẫu số liệu ghép nhóm là

.

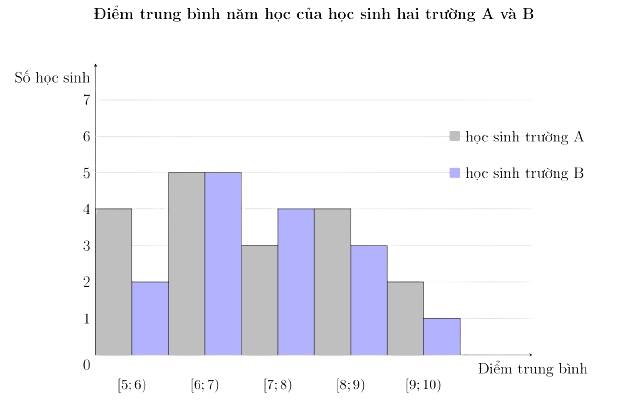
Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm là .

**⑵** Hỏi chiều cao của cây dừa này có phải là giá trị ngoại lệ không?

Vì  và  nên cây dừa có chiều cao 8,4 m là giá trị ngoại lệ của mẫu số liệu ghép nhóm.

**Ví dụ 4.2.**

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



**⑴** Hãy xác định giá trị đại diện cho mỗi nhóm và lập bảng tần số ghép nhóm cho mẫu số liệu trên.

**⑵** 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?



🖎 ***Lời giải***

**⑴** Hãy xác định giá trị đại diện cho mỗi nhóm và lập bảng tần số ghép nhóm cho mẫu số liệu trên.

Ta có:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Điểm trung bình |  |  |  |  |  |
| Giá trị đại diện | 5,5 | 6,5 | 7,5 | 8,5 | 9,5 |
| Học sinh trường A | 4 | 5 | 3 | 4 | 2 |
| Học sinh trường B | 2 | 5 | 4 | 3 | 1 |

**⑵** 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?

▪***Xét mẫu số liệu trường A.***

Cỡ mẫu .

Gọi  là mẫu số liệu gốc gồm điểm trung bình của 18 học sinh trường A được sắp xếp theo thứ tự không giảm.

Ta có: .

Tứ phân vị thứ nhất của mẫu số liệu gốc là .

Do đó, tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là:

.

Tứ phân vị thứ ba của mẫu số liệu gốc là .

Do đó, tứ phân vị thứ ba của mẫu số liệu ghép nhóm là:

.

Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm là: .

▪***Xét mẫu số liệu ở trường B.***

Cỡ mẫu .

Gọi  là mẫu số liệu gốc gồm điểm trung bình của 15 học sinh trường B được sắp xếp theo thứ tự không giảm.

Ta có: .

Tứ phân vị thứ nhất của mẫu số liệu gốc là .

Do đó, tứ phân vị thứ nhất của mẫu số liệu ghép nhóm là:

.

Tứ phân vị thứ ba của mẫu số liệu gốc là .

Do đó, tứ phân thứ ba của mẫu số liệu ghép nhóm là:

.

Vậy khoảng tứ phân vị của mẫu số liệu ghép nhóm là: .

Nếu so sánh theo khoảng tứ phân vị thì học sinh trường B có điểm trung bình đồng đều hơn.