

	VIETTEL AI RACE	Public 509
	BIỂU ĐỒ PHÂN TÁN	Lần ban hành: 1

Biểu đồ phân tán là một dạng hình vẽ, biểu thị mối quan hệ giữa hai thông số nhất định nào đó, xác định xem chúng có mối quan hệ với nhau không (còn được gọi là đồ thị X-Y). Ví dụ: quan hệ giữa tỉ lệ khuyết tật sản phẩm và thông số nhiệt độ.

1. Mục đích, ý nghĩa và lợi ích áp dụng

Biểu đồ phân tán được sử dụng để giải quyết các vấn đề và xác định điều kiện tối ưu bằng cách phân tích định lượng mối quan hệ nhân quả giữa hai nhân tố. Biểu đồ phân tán được sử dụng để:

- Xác định mức độ tương quan giữa hai biến số;
- Xác định có tồn tại mối quan hệ giữa hai đặc tính bằng cách đánh dấu các cặp số liệu trên hệ toạ độ X-Y hoặc đánh dấu một đặc tính trên trục Y còn đặc tính khác trên trục X;

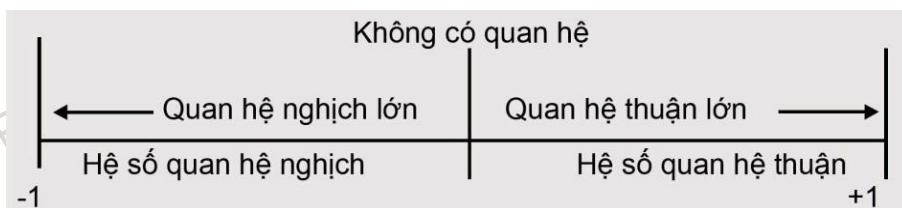
Khi xác định có tồn tại mối quan hệ giữa hai đặc tính tức là chúng có quan hệ với nhau. Khi một đặc tính tăng thì đặc tính khác cũng tăng, chúng có mối quan hệ thuận và các điểm dữ liệu sẽ nằm trong vùng elip nghiêng về bên phải. Nếu một đặc tính giảm mà đặc tính khác tăng chúng có mối quan hệ nghịch. Khi đó, các điểm của dữ liệu nằm trong vùng elip

	VIETTEL AI RACE BIỂU ĐỒ PHÂN TÁN	Public 509 Lần ban hành: 1
--	---	----------------------------------

nghiêng về bên trái.

Khi xác định hai đặc tính không có mối quan hệ, các điểm dữ liệu phân tán trong một vòng tròn.

Mối quan hệ được thể hiện qua giá trị của hệ số hồi qui hay hệ số tương quan r . Giá trị này gần với -1 (<-0.85) thì có mối quan hệ nghịch rất lớn (quan hệ nghịch chật). Giá trị gần với $+1$ (>0.85) thì có mối quan hệ thuận rất lớn (quan hệ thuận chật). Giá trị này gần 0 thì mối quan hệ giữa hai đặc tính rất kém.



Hình 1.5: Hình biểu diễn mức độ quan hệ của 2 biến số theo hệ số tương quan

Trên cơ sở phân tích biểu đồ phân tán, hệ số tương quan r và phương trình hồi qui, có thể dự báo được đặc tính chất lượng và đặc tính quá trình, cũng như xác định các yếu tố cần được kiểm soát chặt chẽ để đảm bảo chất lượng sản phẩm.

Dựa vào việc phân tích biểu đồ có thể thấy được

	VIETTEL AI RACE	Public 509
	BIỂU ĐỒ PHÂN TÁN	Lần ban hành: 1

nhân tố này phụ thuộc như thế nào vào một nhân tố khác và mức độ phụ thuộc giữa chúng.

2. Phân tích biểu đồ phân tán

Mỗi quan hệ giữa các đặc tính nghĩa là sự thay đổi của một đặc tính có khả năng làm thay đổi các đặc tính khác. Nguyên tắc của loại biểu đồ này là phân tích mối liên hệ giữa hai đặc tính (biến số). Mô hình chung của loại biểu đồ này gồm:

- Trục hoành dùng để biểu thị những biến số.
- Trục tung dùng để biểu thị số lượng biến số hay tần số.
- Hình dạng của biểu đồ có thể là những đám chấm, đường gấp khúc hay đường vòng.

Để phân tích biểu đồ phân tán chúng ta có thể:

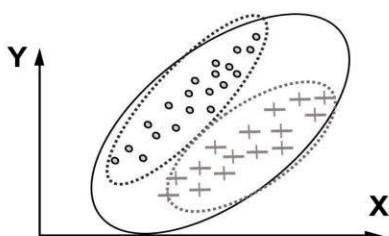
- Dựa vào hình dạng biểu đồ để dự đoán mối tương quan;
- Tính toán hệ số tương quan để có thể khẳng định chắc chắn hơn về mối tương quan đã nhận định dựa trên hình dạng của biểu đồ.



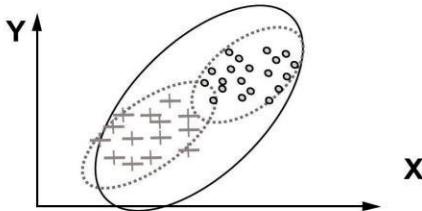
ST T	DẠNG BIỂU ĐỒ	MÔI TƯƠNG QUAN
1		Môi tương quan thuận (đồng biến) $X \nearrow \rightarrow Y \nearrow$ $X \searrow \rightarrow Y \searrow$
2		Có thể xuất hiện môi tương quan thuận để khẳng định ta phải tính hệ số tương quan r
3		Môi tương quan nghịch $X \nearrow \rightarrow Y \searrow$ $X \searrow \rightarrow Y \nearrow$

4		Có thể xuất hiện mối tương quan nghịch giống như trường hợp 2 để khẳng định phải tính hệ số tương quan r
5		Không có tương quan
6		Mối tương quan không tuyến tính

Lưu ý: Khi phân tích tương quan chúng ta phải xét đến các vấn đề như Cách chia trực toạ độ; Sự phân vùng dữ liệu; Những mối tương quan giả.



Hình 1.6: Có mối tương quan khi phân vùng dữ liệu



Hình 1.7: Không có mối tương quan khi phân vùng dữ liệu

3. Phân tích hệ số tương quan

Để mô tả độ tương quan giữa hai biến cần phải ước tính hệ số tương quan r . Công thức hệ số tương quan như sau:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{\sum_{i=1}^n x_i y_i - \frac{1}{n} \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)}{(n - 1) s_x s_y}$$

Công thức này còn được biết đến như là hệ số tương quan Pearson (Pearson's correlation coefficient) để ghi nhận công hiến của nhà thống kê học nổi tiếng Karl Pearson. Nếu giá trị của r là dương, hai biến x và y cùng biến thiên theo một hướng. Nếu giá trị của r là âm, x và y có mối quan hệ nghịch, tức khi x tăng thì y

	VIETTEL AI RACE	Public 509
	BIỂU ĐỒ PHÂN TÁN	Lần ban hành: 1

giảm và ngược lại. Nếu $r = 1$ hay $r = -1$ như Biểu đồ 1.8 a và 1.8 b, mối liên hệ của y và x được hoàn toàn xác định, có nghĩa là cho bất cứ giá trị nào của x , chúng ta có thể xác định giá trị của y . Nếu $r = 0$ như biểu đồ 1.8 c, hai biến x và y hoàn toàn độc lập, tức không có liên hệ với nhau.

2025-10-19 02.59.58_AI Race