

# Xác suất và thống kê

Nguyễn Tấn Nhựt

Ngày 27 tháng 9 năm 2024

Đối tượng nghiên cứu của lý thuyết xác suất là các hình mẫu ngẫu nhiên có hai tính chất:

**Không chắc chắn về kết quả:** Chúng ta không biết chính xác kết quả trước khi sự kiện xảy ra tạo thành kết cục nhưng có thể mô tả bằng các khả năng xảy ra khác nhau.

**Tính qui luật trong dài hạn:** Mặc dù kết quả của một lần xảy ra có vẻ ngẫu nhiên, nhưng khi sự kiện được lặp đi lặp lại nhiều lần, sẽ có một mô hình hay xu hướng có thể quan sát được. Xác suất giúp mô tả xu hướng đó.

Trong đời sống, chúng ta thường xuyên đối mặt với những tình huống mang tính ngẫu nhiên làm cho chúng ta không thể quyết định ...

Ngẫu nhiên dẫn đến bất định.

Trong toán học và khoa học, có hai công cụ chủ chốt giúp chúng ta đối phó với sự bất định này: Xác suất và Thống kê. Tuy khác nhau về cách tiếp cận, cả hai đều giúp chúng ta hiểu rõ hơn về những gì có thể xảy ra và suy luận về những hiện tượng mà ta chưa thể chắc chắn.

## 1 Xác suất: Dự đoán sự ngẫu nhiên

Hãy bắt đầu với xác suất. Xác suất là công cụ toán học dùng để đo lường khả năng xảy ra các sự kiện. Nếu chúng ta có một đồng xu công bằng - nghĩa là đồng xu không thiên vị mặt nào - thì xác suất xuất hiện mặt sấp hay mặt ngửa đều bằng nhau, tức 50%. Điều này có nghĩa rằng, nếu ta gieo đồng xu 100 lần, lý thuyết xác suất dự đoán rằng số lần xuất hiện mặt sấp và mặt ngửa sẽ gần như bằng nhau.

Tại sao vậy? Xác suất dựa trên các mô hình lý thuyết để đưa ra dự đoán. Trong trường hợp này, mô hình xác định rằng mỗi lần gieo đồng xu, có 50% khả năng là mặt sấp và 50% khả năng là mặt ngửa. Do đó, khi ta lặp lại thí nghiệm nhiều lần, kết quả dự kiến là số lần xuất hiện mỗi mặt sẽ cân bằng nhau. Đây là cách mà xác suất giúp chúng ta hiểu dự đoán các hiện tượng ngẫu nhiên mà ta chưa có dữ liệu thực tế.

## 2 Thống kê: Suy luận từ dữ liệu

Ngược lại với xác suất, thống kê bắt đầu bằng cách thu thập dữ liệu thực tế. Giả sử bạn thực hiện một thử nghiệm bằng cách gieo đồng xu 100 lần và quan sát kết quả. Nếu số lần xuất hiện mặt sấp và mặt ngửa gần bằng nhau, chẳng hạn như 51 lần sấp và 49 lần ngửa, bạn có thể suy luận rằng đồng xu có khả năng xuất hiện mặt sấp và mặt ngửa bằng nhau. Điều này phản ánh một đặc điểm quan trọng của thống kê: nó cho phép chúng ta suy luận ngược từ dữ liệu về tổng thể.

Thống kê là công cụ giúp chúng ta giải thích và đưa ra kết luận dựa trên các quan sát thực tế. Khi bạn có dữ liệu từ một mẫu nhỏ (ở đây là 100 lần gieo đồng xu), thống kê cho phép bạn ước lượng đặc điểm của tổng thể lớn hơn (đồng xu là công bằng hay không). Sự khác biệt giữa thống

kê và xác suất ở đây là thống kê dựa vào kết quả thực tế, trong khi xác suất dựa trên mô hình lý thuyết.

### 3 So sánh giữa xác suất và thống kê

Dù cả hai đều xử lý sự ngẫu nhiên và bất định, xác suất và thống kê có cách tiếp cận rất khác nhau:

- **Xác suất:** Xác suất bắt đầu với việc mô hình hóa tổng thể hoặc sự kiện, từ đó dự đoán các kết quả có thể xảy ra. Khi ta biết đồng xu công bằng, xác suất cho phép ta dự đoán rằng nếu gieo đồng xu 100 lần, số lần xuất hiện mặt sấp và mặt ngửa sẽ gần như bằng nhau.
- **Thống kê:** Thống kê bắt đầu từ việc thu thập dữ liệu từ các thử nghiệm thực tế, sau đó đưa ra suy luận về tổng thể. Nếu ta không biết tính chất của đồng xu và chỉ quan sát 100 lần gieo, thì thống kê giúp ta kết luận rằng đồng xu xó thể là công bằng nếu số lần xuất hiện mặt sấp và mặt ngửa gần như bằng nhau.

.

### 4 Xác suất và thống kê

Xác suất và thống kê có mối liên hệ chặt chẽ với nhau vì tất cả các phát biểu thống kê đều là các phát biểu về xác suất. Mặc dù vậy, đôi khi hai khái niệm này lại có vẻ như là những chủ đề rất khác nhau. Xác suất là một phạm trù logic độc lập; có một số quy tắc và câu trả lời đều tuân theo logic từ các quy tắc, mặc dù việc tính toán có thể phức tạp. Trong thống kê, chúng ta áp dụng xác suất để rút ra kết luận từ dữ liệu. Điều này có thể lộn xộn và thường liên quan đến nghệ thuật cũng như khoa học.

Ví dụ về xác suất: Bạn có một đồng xu công bằng (xác suất mặt sấp hoặc mặt ngửa bằng nhau). Bạn sẽ tung nó 100 lần. Xác suất xuất hiện 60 mặt ngửa trở lên là bao nhiêu? Chỉ có một câu trả lời (khoảng 0,028444) và chúng ta sẽ học cách tính toán.

Ví dụ về thống kê: Bạn có một đồng xu không rõ nguồn gốc. Để xem xét xem nó có công bằng không, bạn tung nó 100 lần và đếm số mặt ngửa. Giả sử bạn đếm được 60 mặt ngửa. Nhiệm vụ của bạn với tư cách là một nhà thống kê là rút ra kết luận (suy luận) từ dữ liệu này. Có nhiều cách để tiến hành, cả về hình thức kết luận và các phép tính xác suất được sử dụng để biện minh cho kết luận. Trên thực tế, các nhà thống kê khác nhau có thể đưa ra các kết luận khác nhau.

---

Xác suất và thống kê thường đi chung với nhau vì chúng hỗ trợ lẫn nhau trong việc phân tích và giải thích dữ liệu. Cụ thể:

1. **Xác suất:** Là nền tảng lý thuyết của thống kê, cung cấp các công cụ toán học để mô tả và đo lường sự không chắc chắn. Xác suất giúp mô hình hóa các hiện tượng ngẫu nhiên, chẳng hạn như việc dự đoán kết quả của các sự kiện không thể đoán trước chính xác nhưng có thể gán xác suất cho chúng. Điều này cung cấp cơ sở cho các phương pháp thống kê.

2. **Thống kê:** Là quá trình thu thập, phân tích, và diễn giải dữ liệu. Thống kê sử dụng các công cụ và mô hình xác suất để đưa ra suy luận và dự đoán dựa trên dữ liệu. Thống kê mô tả dữ liệu thông qua các chỉ số như trung bình, phương sai, và xác suất giúp ước lượng hoặc kiểm định các giả thuyết về dữ liệu.

Kết hợp lại:

Xác suất: giúp xác định mô hình lý thuyết, từ đó hiểu rõ các mẫu ngẫu nhiên và đưa ra dự đoán về kết quả trong tương lai.

Thống kê: sử dụng các công cụ xác suất để ước lượng và kiểm định các giả thuyết từ dữ liệu thực tế, qua đó kiểm chứng hoặc điều chỉnh mô hình xác suất đã xây dựng.

Vì vậy, xác suất là lý thuyết cơ bản cho thống kê, còn thống kê là ứng dụng của lý thuyết xác suất để phân tích dữ liệu và ra quyết định.