1> Nghiên cứu và ứng dụng các hàm phân bố với biễn ngẫu nhiên rời rạc

2> Nghiên cứu và ứng dụng các hàm phân bố với biên ngẫu nhiên liên tục

3> Tìm hiểu dãy Markov và ứng dụng của nó

Câu 1:

- Phân phối Bernoulli: Một phép thử Bernoulli có kết quả nhận được là một trong hai giá trị hoặc "thành công" hoặc "thất bại". "Thành công" xảy ra với xác suất là p, "thất bại" với xác suất là q=1-p. Tham số p là số thực nằm giữa 0 và 1. Một biến ngẫu nhiên X có phân phối Bernoulli nhận một  trong 2 giá trị: 1 (thành công) hoặc 0 (thất bại). Xác suất thành công , và xác suất thất bại

- Phân phối nhị thức: Phân phối nhị thức với tham số p và n là tổng của n phép thử Bernoulli với xác suất p độc lập với nhau. Biến ngẫu nhiên có phân phối nhị thức nhận giá trị từ 0 đến n và xác suất để chọn ra x phần tử mong muốn trong n phần tử là  với x=0,1,2,...n.

- Phân phối poison: Trong lý thuyết xác suất và thống kê, phân phối Poisson là một phân phối xác suất rời rạc. Nó khác với các phân phối xác suất rời rạc khác ở chỗ thông tin biết không phải là xác suất để một sự kiện xảy ra thành công trong một lần thử như trong phân phối Bernoulli hay là số lần mà sự kiện đó xảy ra trong n lần thử như trong phân phối nhị thức, mà chính là trung bình số lần xảy ra thành công của một sự kiện trong một khoảng thời gian nhất định. Gía trị trung bình này được gọi là lambda, kí hiệu là . Phân phối Poisson còn được dùng cho khoảng mà đơn vị khác thời gian như: khoảng cách, diện tích hay thể tích. Một ví dụ cổ điển là sự phân rã hạt nhân của các nguyên tử. Khi những sự kiện xảy ra một cách ngẫu nhiên đều đặn với tỷ lệ là   sự kiện trên một đơn vị thời gian, khi đó biến ngẫu nhiên X đếm số sự kiện xảy ra trong khoảng thời gian độ dài t, có phân phối Poisson.

+ Ứng dụng: Có rất nhiều lĩnh vực mà việc áp dụng định lý nhị thức là không thể tránh khỏi, ngay cả trong các lĩnh vực thế giới hiện đại như máy tính. Trong các lĩnh vực máy tính, định lý nhị thức rất hữu ích trong việc phân phối địa chỉ IP. Với định lý nhị thức, phân phối tự động của địa chỉ IP không chỉ có thể thực hiện được mà còn là phân phối của địa chỉ IP ảo.

Câu 2:

Phân phối chuẩn (Normal Distribution) là sự phân bố dữ liệu mà ở đó giá trị tập trung nhiều nhất ở khoảng giữa và các giá trị còn lại rải đều đối xứng về phía các điểm cực trị. Chiều cao là một ví dụ đơn giản của trường hợp tuân theo phân phối chuẩn: hầu hết mọi người có chiều cao trung bình, số người cao hơn hoặc thấp hơn chiều cao trung bình tương đối đồng đều và chỉ có một số rất nhỏ (nhưng vẫn gần như tương đương) số lượng người rất cao hoặc rất thấp.

Biểu diễn đồ thị của một phân phối chuẩn đôi khi được gọi là đường cong hình chuông vì hình dạng loe rộng ra của nó. Hình dạng chính xác có thể thay đổi tùy theo tập toàn thể của phân phối nhưng đỉnh luôn luôn ở giữa và đường cong luôn đối xứng. Trong một phân phối chuẩn, giá trị trung bình, yếu vị và trung vị là giống nhau.

Trong xác suất, phân phối chuẩn được ứng dụng trong việc tính xác suất các biến ngẫu nhiên liên tục như chiều cao, cân nặng, huyết áp, mật độ xương..

Câu 3:

**Định nghĩa**

Một quá trình Markov là một quá trình ngẫu nhiên thỏa mãn tính chất Markov (đôi khi được gọi là “tính không ghi nhớ”). Nói đơn giản, nó là một quá trình mà các kết quả ở tương lai có thể được dự đoán chỉ dựa trên trạng thái hiện tại và—quan trọng hơn—dự đoán ấy tốt bằng dự đoán dựa trên toàn bộ lịch sử của quá trình đó. Nói cách khác, dựa trên trạng thái hiện tại của hệ thống, những trạng thái quá khứ và tương lai là độc lập.

Một xích Markov là một loại quá trình Markov có không gian trạng thái rời rạc hoặc tập chỉ số rời rạc (thường biểu diễn thời gian), tuy nhiên không có định nghĩa chính xác thống nhất. Thông thường, một xích Markov còn được định nghĩa là một quá trình Markov trong thời gian liên tục hoặc rời rạc với không gian trạng thái đếm được (tức thời gian bất kỳ), nhưng cũng có định nghĩa khác coi xích Markov có thời gian rời rạc trong không gian trạng thái đếm được hoặc liên tục (tức không gian trạng thái bất kỳ).

**Các loại xích Markov**

Không gian trạng thái và hệ số thời gian của hệ thống cần được xác định rõ ràng. Bảng sau đưa ra cái nhìn tổng quan về các loại quá trình Markov khác nhau dựa trên các mức độ không gian trạng thái và thời gian rời rạc hoặc liên tục:

|  |  |  |
| --- | --- | --- |
|  | **KHÔNG GIAN TRẠNG THÁI ĐẾM ĐƯỢC** | **KHÔNG GIAN TRẠNG THÁI LIÊN TỤC HOẶC TỔNG QUÁT** |
| **THỜI GIAN RỜI RẠC** | Xích Markov thời gian rời rạc trên một không gian trạng thái hữu hạn hoặc đếm được | Xích Markov trên không gian trạng thái đo được (ví dụ như xích Harris) |
| **THỜI GIAN LIÊN TỤC** | Quá trình Markov thời gian liên tục | Bất kỳ quá trình ngẫu nhiên liên tục nào với tính chất Markov (ví dụ như quá trình Wiener) |

Để ý rằng không có sự thống nhất trong việc dùng các thuật ngữ để chỉ trường hợp đặc biệc của quá trình Markov. Thường cụm từ “xích Markov” dùng để chỉ xích Markov thời gian rời rạc (DTMC), nhưng một số tác giả dùng “quá trình Markov” để chỉ một xích Markov thời gian liên tục (CTMC) mà không nói cụ thể. Ngoài ra, có một số mở rộng của quá trình Markov nằm ngoài bốn loại trên nhưng không có tên gọi đặc biệt nào (xem mô hình Markov). Hơn nữa, chỉ số thời gian không nhất thiết phải là số thực; giống như không gian trạng thái, có những quá trình di chuyển qua tập chỉ số của những cấu trúc toán học khác.

Trong khi tham số thời gian thường là rời rạc, không gian trạng thái của một chuỗi Markov không có quy tắc thống nhất chung nào. Tuy nhiên, nhiều ứng dụng của chuỗi Markov thường dùng không gian trạng thái hữu hạn hoặc vô hạn đếm được, cho phép phân tích thống kê đơn giản hơn. Ngoài số chỉ thời gian và không gian trạng thái, còn có nhiều biến thể, mở rộng và tổng quát khác. Để đơn giản, phần lớn bài viết tập trung vào trường hợp thời gian rời rạc, không gian trạng thái rời rạc, trừ khi được nói rõ.

**Chuyển đổi**

Sự biến đổi trạng thái của hệ thống được gọi là sự **chuyển trạng thái** hay **chuyển đổi** (tiếng Anh: *transition*). Xác suất của các khả năng chuyển trạng thái khác nhau được gọi là xác suất chuyển đổi. Một quá trình được xác định bởi một không gian trạng thái, một ma trận xác suất chuyển đổi mô tả xác suất của các chuyển đổi, và một trạng thái ban đầu (hay phân bố ban đầu) trên không gian trạng thái. Theo thông lệ, ta coi tất cả trạng thái và chuyển đổi khả dĩ được tính trong quá trình, nên luôn có một trạng thái tiếp theo, và quá trình không dừng lại.

Một quá trình ngẫu nhiên thời gian rời rạc gồm một hệ thống ở một trạng thái nhất định vào một bước nào đó, và trạng thái của hệ thay đổi giữa các bước. Các bước này thường được coi là khoảnh khắc trong thời gian, nhưng cũng có thể là khoảng cách vật lý hay những đại lượng khác. Về mặt toán học, các bước này được biểu diễn bằng số nguyên hoặc số tự nhiên. Tính chất Markov nói rằng phân bố xác suất có điều kiện của hệ ở bước tiếp theo (và tất cả các bước trong tương lai) chỉ phụ thuộc vào trạng thái hiện tại của hệ, chứ không phụ thuộc vào trạng thái của hệ ở những bước trước.

Do hệ thay đổi ngẫu nhiên, không thể đoán chính xác được trạng thái của xích Markov ở một thời điểm trong tương lai. Tuy nhiên, các tính chất thống kê của hệ có thể được dự đoán.

**Ứng dụng**

Chuỗi markov cũng có nhiều ứng dụng trong mô hình sinh học, đặc biệt là trong tiến trình dân số- một tiến trình tương tự như tiến trình sinh học.

Một ứng dụng của chuỗi Markov gần đây là ở thống kê địa chất. Đó là: chuỗi Markov được sử dụng trong mô phỏng 3 chiều của giá trị có điều kiện riêng phần. Ứng dụng này được gọi là “chuỗi markov thống kê địa chất”, cũng giống như ngành thống kê địa chất. Hiện nay phương pháp này vẫn còn đang phát triển.

Chuỗi Markov cũng có thể ứng dụng trong nhiều trò chơi game. Nhiều loại game của trẻ em (Chutes and Ladders, Candy Land), là kết quả chính xác điển hình của chuỗi Markov. Ở mỗi vòng chơi, người chơi bắt đầu chơi từ trạng thái định sẵn, sau đó phải có lợi thế gì đó mới có thể vượt qua được bàn kế.

Trong ngành quản lý đất đai: người ta còn ứng dụng GIS, RS và chuỗi Markov vào phân tích sự thay đổi sử dụng đất (land use change), từ đó dự báo được tình hình sử dụng đất trong giai đoạn kế tiếp.