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

**Phân phối Bernoulli**

Phân phối Bernoulli là một phân phối xác suất rời rạc của biến ngẫu nhiên chỉ nhận hai giá trị 0 hoặc 1, trong đó gái trị 1 đạt được với xác suất ( gọi là xác suất thành công) và giá trị 0 đạt được với xác suất q = p - 1 (gọi là xác suất thất bại). Nếu là một biến ngẫu nhiên với phân phối này, kí hiệu x ~ Bernoulli(p), ta sẽ có :

*p*(*x* = 1) = *#, p*(*x* = 0) = 1 *# p*(*x* = 1) = 1 *# #*

Một ví dụ cổ điển về biến ngẫu nhiên Bernoulli là kết quả của việc tung một đồng xu (có thể không

đồng chất), mặt chẵn ngửa ứng với giá trị 1, mặt chẵn ứng với giá trị 0. Đồng xu có thể xuất hiện mặt ngữa với xác xuất p và mặt chẵn với xác suất 1 – p.

**Phân phối** Categorical

Cũng là biến ngẫu nhiên rời rạc, nhưng trong hầu hết các trường hợp, đầu ra có thể là một trong nhiều hơn hai giá trị khác nhau. Ví dụ, một bức ảnh có thể chứa một chiếc xe, một người, hoặc một con mèo. Khi đó, ta dùng phân bố tổng quát của Bernoulli distribution và được gọi là Categorical distribution. Các đầu ra được mô tả bởi 1 phần tử trong tập {1,2,…,K}{1,2,…,K}.

Nếu có KK đầu ra có thể đạt được, Categorical distribution sẽ được mô tả bởi KK tham số, viết dưới dạng vector: λ=[λ1,λ2,…,λK]λ=[λ1,λ2,…,λK] với các λkλk không âm và có tổng bằng 1. Mỗi giá trị λkλk thể hiện xác suất để đầu ra nhận giá trị kk:p(x=k)=λkp(x=k)=λkViết gọn lại:p(x)=Catx[λ]p(x)=Catx[λ]

Biểu diễn theo cách khác, ta có thể coi như đầu ra là một vector ở dạng one-hot vector, tức x∈{e1,e2,…,eK}x∈{e1,e2,…,eK} với ekek là vector đơn vị thứ kk, tức tất cả các phần tử bằng 0, trừ phần tử thứ kk bằng 1. Khi đó, ta sẽ có:p(x=ek)=K∏j=1λxjj=λk (31)

**Univariate normal distribution** (Phân phối chuẩn một biến)

Phân phối chuẩn 1 biến (univariate normal hoặc Gaussian distribution) được định nghĩa trên các biến liên tục nhận giá trị x∈(−∞,∞)x∈(−∞,∞).

Phân phối này được mô tả bởi hai tham số: *mean* μμ và *variance* σ2σ2. Giá trị μμ có thể là bất kỳ số thực nào, thể hiện vị trí của *peak*, tức tại đó mà hàm mật độ xác suất đạt giá trị cao nhất. Giá trị σ2σ2 là một giá trị dương, với σσ thể hiện *độ rộng* của phân bố này. σσ lớn chứng tỏ khoảng giá trị đầu ra biến đổi mạnh, và ngược lại.

Hàm mật độ xác suất của phân phối này được định nghĩa là:p(x)=1√2πσ2exp(−(x−μ)22σ2) (32)p(x)=12πσ2exp⁡(−(x−μ)22σ2) (32)Dạng gọn hơn:p(x)=Normx[μ,σ2]

Ví dụ về đồ thị hàm mật độ xác suất của univariate normal distribution được cho trên dưới

Chart, histogram

Description automatically generated

### Multivariate normal distribution

Đây là trường hợp tổng quát của normal distribution khi biến là nhiều chiều, giả sử là DD chiều. Có hai tham số mô tả phân phối này: mean vector μ∈RDμ∈RD và covariance matrix Σ∈SD++Σ∈S++D là một ma trận đối xứng xác định dương.

Hàm mật độ xác suất có dạng:p(x)=1(2π)D/2|Σ|1/2exp(12(x−μ)TΣ−1(x−μ)) (33)p(x)=1(2π)D/2|Σ|1/2exp⁡(12(x−μ)TΣ−1(x−μ)) (33)với |Σ||Σ| là định thức của ma trận hiệp phương sai ΣΣ.

Hoặc viết gọn:p(x)=Normx[μ,Σ]p(x)=Normx[μ,Σ]

Ví dụ về hàm mật độ xác suất của một bivariate normal distribution (2 biến) được cho trên Hình 2b). Các level-sets của mặt này đều là các hình Ellipse đồng tâm.

### 

### Beta distribution

Beta distribution là một phân phối liên tục được định nghĩa trên một biến ngẫu nhiên λ∈[0,1]λ∈[0,1]. Phân phối này phù hợp với miệc mô tả sự biến động (uncertainty) của tham số λλ trong Bernoulli distribution. Nhắc lại, Beta distribution được dùng để mô tả tham số cho một distribution khác. Các bạn sẽ thấy rõ hơn trong một ví dụ ở bài tiếp theo.

Beta distribution được mô tả bởi hai tham số dương α,β∈(0,∞)α,β∈(0,∞). Hàm mật độ xác suất của nó là:p(λ)=Γ(α+β)Γ(α)Γ(β)λα−1(1−λ)β−1 (34)p(λ)=Γ(α+β)Γ(α)Γ(β)λα−1(1−λ)β−1 (34)

với Γ(.)Γ(.) là gamma function:Γ(z)=∫∞0tz−1exp(−t)dtΓ(z)=∫0∞tz−1exp⁡(−t)dtvà có liên quan tới giai thừa khi zz là số tự nhiên:Γ[z]=(z−1)!

### Dirichlet distribution

Dirichlet distribution chính là trưởng hợp tổng quát của Beta distribution khi được dùng để mô tả tham số của Categorical distribution (nhắc lại rằng Categorical distribution là trường hợp tổng quát của Bernoulli distribution).

Dirichlet distribution được định nghĩa trên KK biến liên tục λ1,…,λKλ1,…,λK trong đó các λkλk không âm và có tổng bằng 1. Bởi vậy, nó phù hợp để mô tả tham số của Categorical distribution.

Có KK tham số dương để mô tả một Dirichlet distribution: α1,…,αKα1,…,αK.

Hàm mật độ xác suất:p(λ1,…,λK)=Γ(∑Kk=1αk)∏Kk=1Γ(αk)K∏k=1λαk−1k (35)p(λ1,…,λK)=Γ(∑k=1Kαk)∏k=1KΓ(αk)∏k=1Kλkαk−1 (35)

Viết gọn:p(λ1,…,λK)=Dirλ1,…,λK[α1,…,αK]

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

**Biến ngẫu nhiên liên tục**

Định nghĩa: Biến ngẫu nhiên X được gọi là liên tục nếu những phân bố xác suất của nó có đạo hàm, trong trường hợp này ta gọi f(x) = F’(x), x ∈ R là hàm mật độ xác suất của biến ngẫu nhiên thì X.

**Các hàm phân bổ với biến ngẫu nhiên liên tục**

**Hàm phân phối xác suất**

Định nghĩa: Hàm số F(x) =P(X < x ) , x ∈ D , được gọi là hàm phân phối (hàm phân bố) xác suất của biến ngẫu nhiên X. Nếu X là một biến ngẫu nhiên rời rạc thì hàm phân phối xác suất của X xác định như sau:

F(x) = x ∈ D

Trong đó : pi là xác suất biến ngẫu nhiên x nhận giá trị xi

**Tính chất của hàm phân phối xác suất**

Tính chất 1:

0≤ F(x) ≤ 1 ∀x

Tính chất 2:

Nếu a là giá trị nhỏ nhất có thể có của X và b là giá trị lớn nhất có thể có của X thì:

F(x)=0 với mọi x ≤ a

F(x)=1 với mọi x > b

Tính chất 3:

Hàm phân phối xác suất của biến ngẫu nhiên là một hàm không giảm.

**Hàm mật độ xác suất**

Định nghĩa: Cho biến ngẫu nhiên liên tục X có hàm phân phối xác suất F(x). Nếu tồn tại hàm số f(x) sao cho:

f(x) = F’(x)

thì hàm số f(x) được gọi là hàm mật độ xác suất của biến ngẫu nhiên X. (F(x) phải là hàm khả vi)

**Tính chất của hàm mật độ xác suất**

Tính chất 1:

f(x) ≥ 0 với x ∀ ∈ D

Tính chất 2:

Tính chất 3:

P(a < X < b) =

Tính chất 4:

F(a) =

**Câu 3:** Tìm hiểu dãy Markov và ứng dụng của nó

Trong lý thuyết xác suất và các lĩnh vực liên quan, quá trình Markov (đặt theo tên của nhà toán học người Nga Andrey Markov) là một quá trình ngẫu nhiên thỏa mãn một tính chất đặc biệt, gọi là tính chất Markov (còn gọi là tính mất trí nhớ). Tính chất này giúp dự báo được tương lai chỉ dựa vào trạng thái hiện tại. Điều này cũng có nghĩa trạng thái tương lai và quá khứ là độc lập nhau. Tuy nhiên về sau, quá trình Markov được mở rộng thành Markov bậc cao, trong đó tương lai phụ thuộc vào hiện tại và một quãng thời gian nào đó trong quá khứ.

Xích Markov là quá trình Markov đặc biệt mà trong đó hoặc có trạng thái rời rạc hoặc thời gian rời rạc. Quá trình Markov được nhà toán học Markov bắt đầu nghiên cứu từ khoảng đầu thế kỷ 20 mặc dù có nhiều nghiên cứu hàng trăm năm trước đó về quá trình này nhưng dưới dạng các biến ngẫu nhiên phụ thuộc. Hai ví dụ quan trọng nhất của quá trình Markov là quá trình Wiener (hay chuyển động Brownian) và quá 13 trình Poisson. Hai quá trình này được coi là quan trọng nhất và là trung tâm của lý thuyết quá trình ngẫu nhiên.

Xích Markov có rất nhiều ứng dụng với vai trò là các *mô hình xác suất* trong các quá trình thực tế. Thuật toán được biết đến là PageRank được thực hiện khởi nguồn cho công cụ tìm kiếm của Google được dựa trên xích Markov.