TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN TOÁN ỨNG DỤNG VÀ TIN HỌC

o0o

ĐỒ ÁN I

TÍNH TOÁN BAYES CHO MÔ HÌNH PHÂN PHỐI HÌNH HỌC

Chuyên ngành: Toán ứng dụng

Người hướng dẫn:	TS. Đỗ Văn Cường
Sinh viên thực hiện:	Ngô Quang Tùng
MSSV:	20196006
Lớp:	HTTTQL - K64

HÀ NỘI, 07/2023

NHẬN XÉT CỦA GIẢNG VIÊN

1. Mục tiêu

(a)

(b)

(c)

2. Nội dung

(a)

(b)

(c)

3. Đánh giá kết quả đạt được

(a)

(b)

(c)

Hà Nội, ngày ... tháng ... năm 2023

Giảng viên hướng dẫn

TS. ĐỖ VĂN CƯỜNG

Mục lục

MỞ ĐẦU

Suy luận Bayes là một phương pháp suy luận thống kê trong đó định lý Bayes được sử dụng để cập nhật xác suất cho một giả thuyết khi có thêm bằng chứng hoặc thông tin. Phương pháp suy luận Bayes là một phương pháp quan trọng trong thống kê toán học và được ứng dụng trong nhiều lĩnh vực như khoa học, kỹ thuật, triết học, y học...
Suy luận Bayes lấy ra xác suất hậu nghiệm là hệ quả của hai tiền đề: xác suất tiên nghiệm và "hàm khả năng" xuất phát từ mô hình thống kê cho dữ liệu quan sát. Ở nội dung trong bài báo cáo này chúng ta sẽ sử dụng suy luận Bayes cho mục đích chính là tính toán Bayes cho mô hình phân phối hình học. Để giải quyết cho vấn đề này chúng ta sẽ được giới qua về cơ sở lý thuyết, cách xây dựng mô hình phân phối hình học bằng phương pháp Bayes, và áp dụng mô hình vào việc ứng dụng, mô phỏng, giải quyết bài toán thực tế.

QUY TẮC VIẾT BÁO CÁO VÀ KÝ HIỆU

Trong báo cáo này có sử dụng một số tên viết như sau:

MLE	Maximum Likelihood Estimation (Ước lượng hợp lý cực đại)

Đồng thời, một số ký hiệu trong báo cáo như:

$\hat{\mu}_{\text{MLE}}$	Giá trị ước lượng hợp lý cực đại
$\hat{\theta}_{Bayes}$	Giá trị ước lượng Bayes
$\underline{x}$	$(x_1, x_2, ..x_n)$
$f(\underline{x}\|\theta)$	Hàm mật độ xác suất đồng thời của các biến ngẫu nhiên cùng phân phối $(X_1, X_2, ...X_n)$
$\pi(\theta)$	Phân phối tiên nnghiệm của θ (prior)
$p(\theta\|\underline{x})$	Phân phối hậu nghiệm của θ (posterior)

CẤU TRÚC CỦA BÁO CÁO

Báo cáo của đồ án I này sẽ được trình bày gồm 4 chương sau:

- **Chương 1: Kiến thức chuẩn bị** trình bày về phân phối hình học, định lý Bayes để sử dụng cho các chương sau.
- **Chương 2: Tính toán Bayes cho mô hình phân phối hình học** chương này sẽ trình bày về cách xây dựng mô hình phân phối hình học cho phương pháp Bayes.
- **Chương 3: Ứng dụng** trình bày mô phỏng dựa trên số liệu mô phỏng và số liệu thực tế.
- **Chương 4: Kết luận và hướng phát triển đề tài**

Dù đã cố gắng xong vẫn không thể tránh khỏi những hạn chế cần khắc phuc. Vì vậy, em rất mong quý thầy cô đưa ra những ý kiến góp ý để đồ án có thể phát triển và có những kết quả tốt hơn.

LỜI CẢM ƠN

Báo cáo này được thực hiện và hoàn thành tại Trường Đại Học Bách Khoa Hà Nội, kì học 2022-2.
Em xin được dành lời cảm ơn tới TS. Đỗ Văn Cường, giảng viên hướng dẫn và là người gợi ý cho đề tài này. Thầy đã hướng dẫn, đưa ra những góp ý bổ ích để em có thể hiểu rõ hơn và hoàn thành đề tài này. Sau cùng, báo cáo này có thể không tránh được những thiếu sót, mong được thầy cô góp ý.

Hà Nội, ngày ... tháng ... năm 2023
Sinh viên thực hiên

Ngô Quang Tùng

1 Kiến thức chuẩn bị

1.1 Phân phối nhị thức âm

Định nghĩa 1.1.1 Phân phối nhị thức âm (Negative binomial distribution) là một phân phối của số lần thành công và thất bại trong một loạt các lần thử độc lập trước khi đạt được một số lần thành công nhất định với hai tham số p và r. Dưới đây là các yếu tố chính cần lưu ý về thí nghiệm nhị thức âm:

- Thí nghiệm cần lặp lại x lần thử.
- Mỗi lần thử sẽ chỉ có hai khả năng xảy ra, một là thành công, hai là thất bại..
- Xác suất thành công là giống nhau ở mỗi lần thử.
- Kết quả của thí nghiệm này độc lập với kết quả của thí nghiệm kia.
- Thí nghiệm cần được thực hiện cho tới khi r lần thành công được ghi nhận với r cho trước.

Xác suất phân phối nhị thức âm có thể được tính như sau:

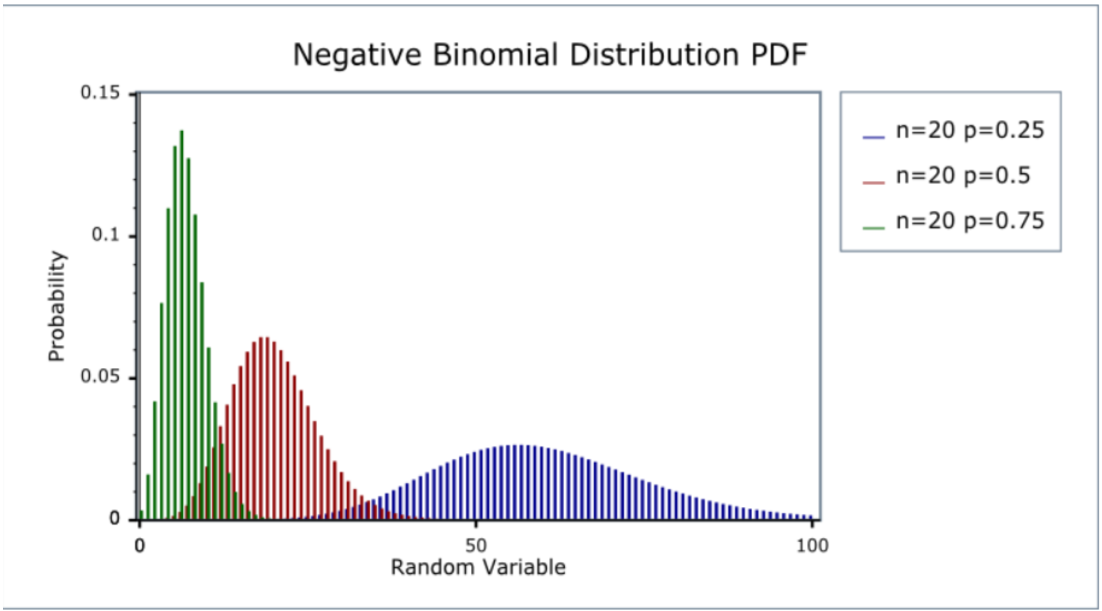
$$P_x(x) = \binom{x-1}{r-1} \times p^r \times (1-p)^{x-r}.$$

Trong đó:

- x = tổng số lần thử.
- r = Tổng số lần thành công mong muốn.
- p = xác suất thành công của mỗi lần thử.
- 1–p = xác suất thất bại của mỗi lần thử.

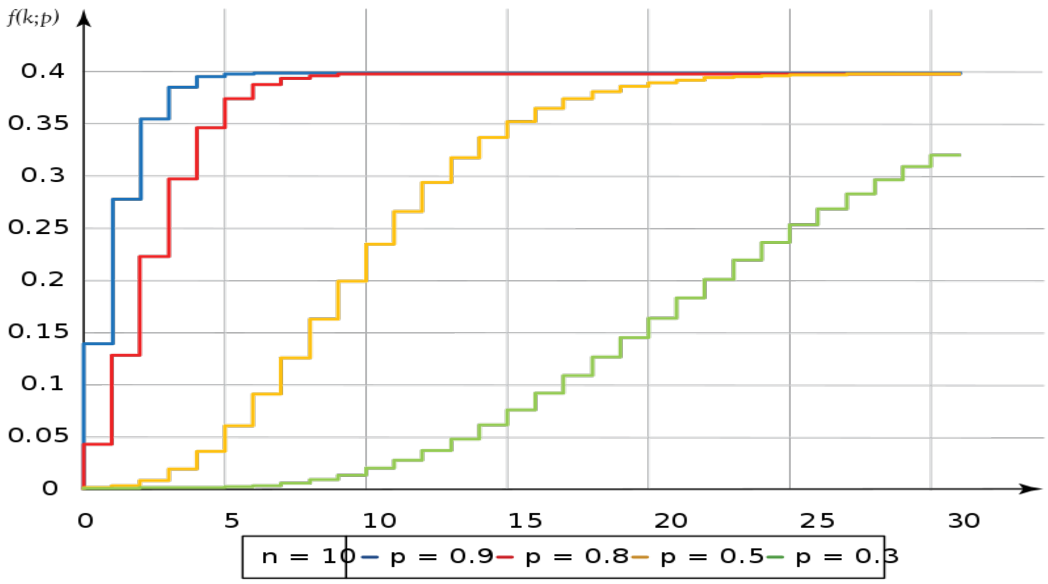
- $P_x(x)$ = xác suất nhị thức âm, xác suất mà một thí nghiệm nhị thức âm n lần thử sẽ đạt được thành công thứ r tại lần thử thứ x với xác suất thành công cho mỗi lần thử là p.
- $\binom{n}{r}$ = Tổ hợp chập n lấy r phần tử.

Hàm khối lượng xác suất



Hình 1: Hàm khối lượng phân phối nhị thức âm với một số tham số p.

Hàm phân phối tích lũy



Hình 2: Hàm phân phối tích lũy nhị thức âm.

1.2 Phân phối hình học

Định nghĩa 1.2.1 Phân phối hình học (Geometric Distribution) là dạng đặc biệt của phân phối nhị thức âm. Nó liên quan tới số lượt thử cần thiết cho một lần thành công duy nhất. Vì vậy phân phối hình học là một phân phối nhị thức âm với số lần thành công (r) là 1.

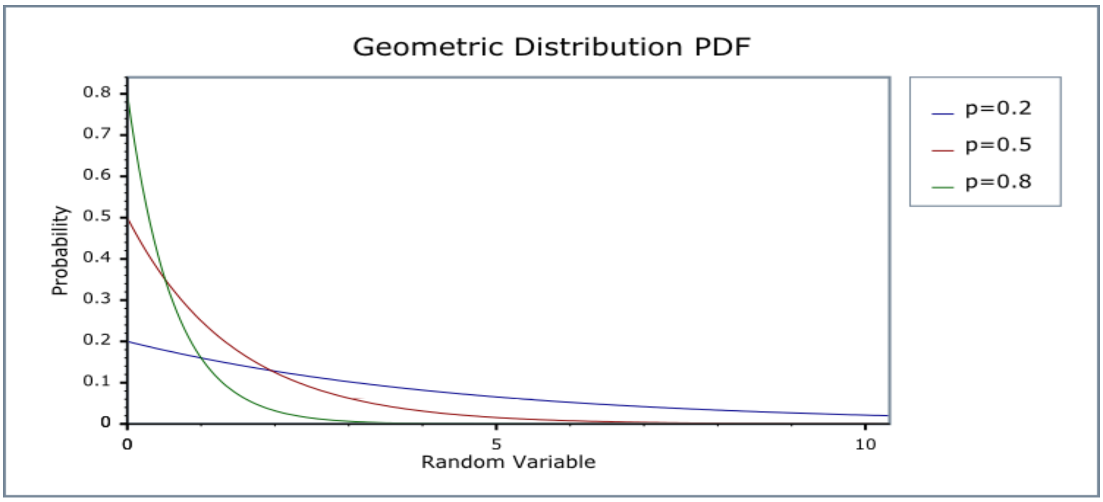
Hàm khối lượng xác suất

Biến ngẫu nhiên rời rạc X được gọi là có phân phối hình học nếu nó có hàm khối lượng xác suất có dạng:

$$P_x(x) = (1-p)^x \times p.$$

trong đó: $0 < \text{p} \leq 1$.
và x $\in \{0, 1, 2, ...\}$
Các đặc trưng của phân phối hình học :

- $E(X) = \frac{1}{p}$
- $V(X) = \frac{1-p}{p^2}$

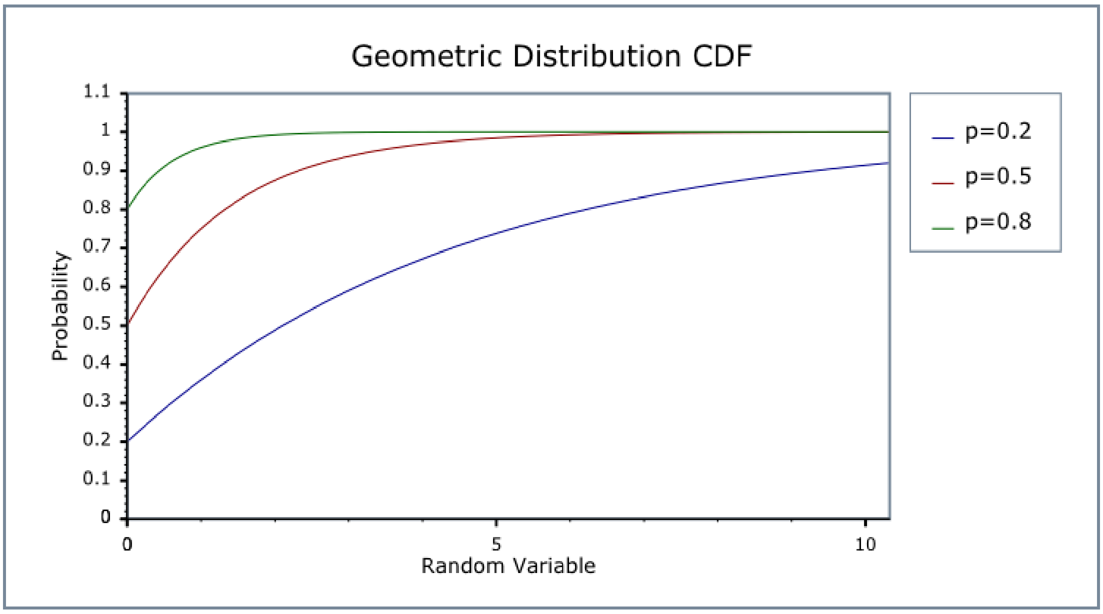


Hình 1: Hàm khối lượng phân phối hình học với một số tham số p.

Hàm phân phối tích lũy

Hàm phân phối tích lũy là xác suất để một biến X có giá trị nhỏ hơn hoặc bằng x, được biểu diễn như sau:

$$P(X \leq x) = 1 - (1-p)^x.$$



Hình 2: Hàm phân phối tích lũy với một số tham số p.

1.3 Phân phối Beta

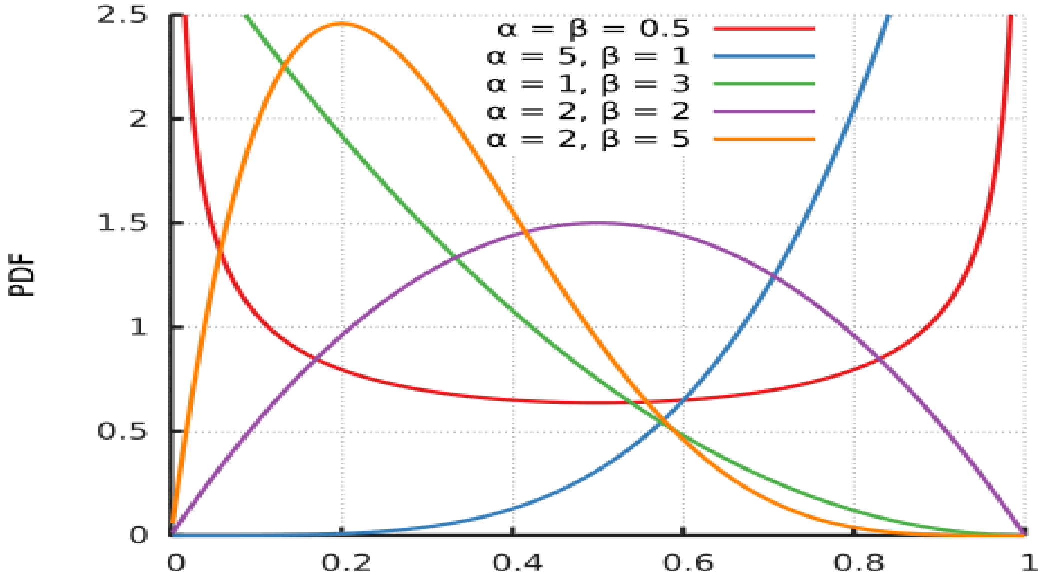
Định nghĩa 1.3.1 Biến ngẫu nhiên liên tục X nhận giá trị trong [0;1] tuân theo phân phối Beta với hai tham số a và b, thường được ký hiệu $X \sim Beta(a, b)$ nên X có hàm mật độ xác suất:

$$P_x(x) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1}.$$

Các đặc trưng của phân phối $Beta$:

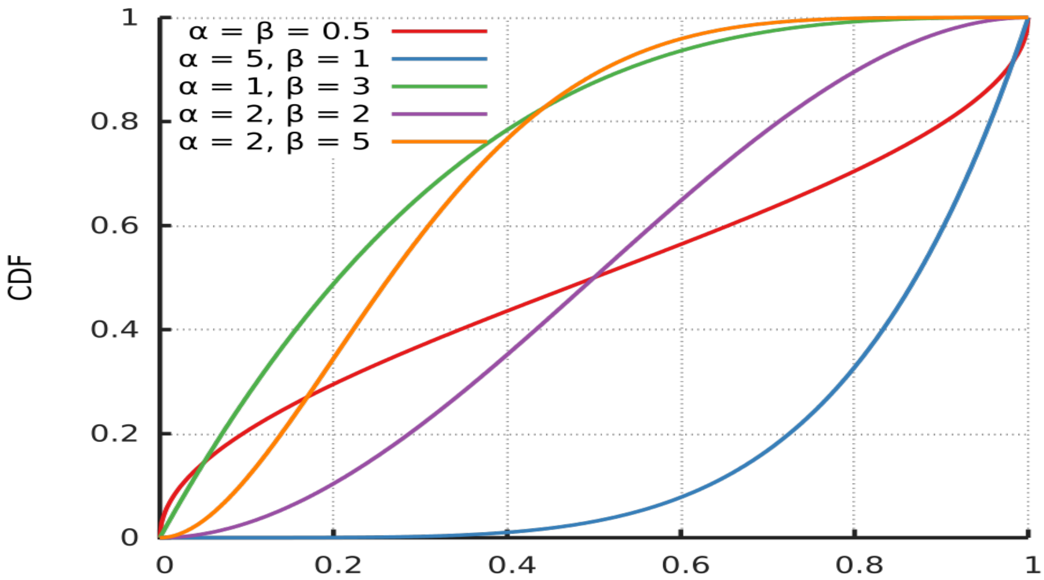
- $E(X) = \frac{a}{a+b}$
- $V(X) = \frac{ab}{(a+b)^2(a+b+1)}$

Hàm mật độ xác suất



Hình 1: Hàm mật độ xác suất của phân phối Beta với một số tham số α và β.

Hàm phân phối tích lũy



Hình 2: Hàm phân phối tích lũy của phân phối Beta với một số tham số α và β.

1.4 Định lý Bayes

Định lý 1.4.1 Định lý Bayes

Gọi $X = \{\underline{x} = (x_1, ..., x_n) | n \in N\}$ là không gian mẫu. Đặt $P = \{P_\theta | \theta \in \Theta\}$ là một mô hình tham số và quan sát D trong không gian mẫu X. Ký hiệu $\pi(\theta)$ là phân phối tiên nghiệm của tham số θ,$f(D|\theta)$ là hàm mật độ xác suất của mô hình, $p(\theta, D)$ là phân phối hậu nghiệm của tham số θ được cho bởi quan sát D. Định lý Bayes phát biểu rằng:

$$p(\theta|D) = \frac{f(D|\theta) \times \pi(\theta)}{\int_\Theta f(D|\theta) \times \pi(\theta) d\theta}$$

Trong đó, $f(D|\theta)$ được gọi là hàm hợp lý (Likelihood), $\int_\theta f(D|\theta) \times \pi(\theta) d\theta$ được gọi là marginal likelihood, là một hằng số, ta thường viết:

$$p(\theta|D) \propto f(D|\theta) \times \pi(\theta)$$
$$\text{posterior} \propto \text{likelihood} \times \text{prior}$$

2 TÍNH TOÁN BAYES CHO MÔ HÌNH PHÂN PHỐI HÌNH HỌC

2.1 Hàm hợp lý và phân phối tiên nghiệm

2.1.1 Phân phối tiên nghiệm

Phân phối tiên nghiệm $\pi(\theta)$ là phân phối xác suất biểu thông tin ban đầu hoặc thông tin cơ bản về tham số θ trước khi quan sát bất kỳ dữ liệu nào. Một số phân phối thường được chọn làm phân phối tiên nghiệm như Gamma, Beta,... Với mô hình này ta sẽ chọn phân phối tiên nghiệm của θ là phân phối Beta hay $\theta \sim Beta(\alpha, \beta)$ với $\alpha > 0$ và $\beta > 0$.
Giả sử trước khi thu thập dữ liệu, ta cho rằng $E(\theta) = \mu$ và $V(\theta) = \sigma^2$. khi đó, ta có thể tính được a,b theo công thức sau:

$$b = \mu - 1 + \frac{\mu \times (1-\mu)^2}{\sigma^2};$$
$$a = \frac{b \times \mu}{1-\mu}.$$

Chứng minh: Ta có:

$$E(\theta) = \frac{a}{a+b} = \mu.$$

Nhân hai vế và rút a theo b ta được:

$$a = a\mu + b\mu,$$
$$a = \frac{b \times \mu}{1-\mu}.$$

Ta lại có:

$$V(\theta) = \frac{ab}{(a+b)^2(a+b+1)} = \sigma^2.$$

Thay a vừa tìm được vào phương trình trên:

$$\frac{\frac{b\times\mu}{1-\mu}\times b}{(\frac{b\times\mu}{1-\mu}+b)^2(\frac{b\times\mu}{1-\mu}+b+1)}=\sigma^2.$$

Rút gọn lại ta được:

$$\frac{\mu\times(1-\mu)^2}{(b+1-\mu)}=\sigma^2.$$

Cuối cùng, nhân hai vế và rút b:

$$b=\mu-1+\frac{\mu\times(1-\mu)^2}{\sigma^2}.$$

Nếu không có thông tin gì về θ, ta thường coi θ có phân phối đều liên tục, tương đương với $\theta \sim Beta(1,1)$. Lúc này phân phối hậu nghiệm chỉ phụ thuộc vào hàm hợp lý.

2.1.2 Hàm hợp lý

Giả sử các biến ngẫu nhiên $X_1, X_2, ...X_n$ độc lập và tuân theo phân phối hình học $X \sim Geo(\theta)$. Khi đó, hàm khối lượng xác suất sẽ là:

$$P_x(x)=(1-p)^x\times p.$$

trong đó: 0 < p $\leq$ 1.
Ta có hàm hợp lý:

$$L(\theta)=\prod_{i=1}^{n}\theta(1-\theta)^{x_i} \ =\theta^n(1-\theta)^{\sum_1^n x_i}.$$

2.1.3 Ước lượng hợp lý cực đại

Mục tiêu của phương pháp ước lượng hợp lý cực đại là đi tìm tham số θ để hàm đạt giá trị lớn nhất.

Định lý 2.1.3 Ước lượng hợp lý cực đại
Đặt $X = \{\underline{x} = (x_1, ..., x_n) | n \in N\}$ là các quan sát độc lập cùng phân phối hình học. Gọi $\hat{\mu}_{\text{MLE}}$ là ước lượng hợp lý cực đại của θ khi đó:

$$\hat{\mu}_{\text{MLE}} = \frac{1}{1+} = \frac{n}{\sum_{i=1}^{n} x_i + n}.$$

Chứng minh: Ta có hàm hợp lý:

$$L(\theta) = \theta^n (1-\theta)^{\sum_1^n x_i}.$$

Ta tìm giá trị cực đại của hàm hợp lý trên cũng tương đương với việc tìm giá trị cực đại hàm Logarit của hàm hợp lý. Do đó, ta có:

$$l(\theta) = lnL(\theta) = nlnp + (\textstyle\sum_1^n x_i) ln(1-\theta).$$

Giải phương trình $\frac{d[lnL(\theta)]}{d\theta} = 0$:

$$\frac{d[lnL(\theta)]}{d\theta} = \frac{n}{\theta} - \frac{\sum_1^n x_i}{1-\theta} = 0.$$

Ta được:

$$\theta = \frac{n}{\sum_{i=1}^{n} x_i + n}.$$

Do đó, ước lượng hợp lý cực đại của tham số θ sẽ là:

$$\hat{\mu}_{\text{MLE}} = \frac{1}{1+\overline{x}} = \frac{n}{\sum_{i=1}^{n} x_i + n}.$$

2.2 Phân phối hậu nghiệm

Tiên nghiệm liên hợp
Từ công thức:

$$\text{Posterior} \propto \text{Likelihood} \times \text{Prior}$$

Giờ ta đã có phân phối tiên nghiệm $\pi(\theta)$ cùng với hàm hợp lý. Khi đó, phân phối hậu nghiệm tỷ lệ với tích của hàm hợp lý và phân phối tiên nghiệm:

$$p(\theta|\underline{x}) \propto f(\underline{x}|\theta) \times \pi(\theta).$$

Ta có hàm hợp lý:

$$f(\underline{x}|\theta) = \theta^n (1-\theta)^{\sum_1^n x_i}.$$

và phân phối tiên nghiệm:

$$\pi(\theta) \propto \theta^{\alpha-1}(1-\theta)^{\beta-1}.$$

Theo định lý Bayes, phân phối hậu nghiệm tỷ lệ với tích của hàm hợp lý và phân phối tiên nghiệm:

$$f(\underline{x}|\theta)\pi(\theta) \propto \theta^n (1-\theta)^{\sum_1^n x_i} \times \theta^{\alpha-1}(1-\theta)^{\beta-1}.$$

Ta có thể viết gọn lại thành:

$$f(\underline{x}|\theta)\pi(\theta) \propto \theta^{\alpha+n-1}(1-\theta)^{\beta+\sum_1^n x_i-1}.$$

Hay nói cách khác phân phối hậu nghiệm là phân phối $Beta(\alpha_1, \beta_1)$
hay $\{\theta|D = d\} \sim Beta(\alpha_1, \beta_1)$. trong đó:

$$\begin{cases} a_1 & = a_0 + n, \\ b_1 & = b_0 + \sum_{i=1}^n x_i. \end{cases}$$

Nhận xét :Khi đặt phân phối tiên nghiệm Beta vào mô hình phân phối hình học khiến phân phối hậu nghiệm cũng là phân phối Beta, ta nói rằng phân phối Beta liên hợp với mô hình phân phối hình học.

2.3 Tính toán hậu nghiệm

2.3.1 Ước lượng Bayes

Với việc ta có phân phối tiên nghiệm là phân phối Beta và phân phối hậu nghiệm $\{\theta|D = d\} \sim Beta(\alpha_1, \beta_1)$, ước lượng Bayes của θ sẽ là:

$$\hat{\theta}_{Bayes} = \frac{\alpha + n}{\alpha + \beta + \sum_{i=1}^{n} x_i + n}.$$

Tiếp theo ta có nhận xét về tính chất tổ hợp lồi của ước lượng hợp lý cực đại và kỳ vọng của phân phối tiên nghiệm.

- Nhắc lại:

$$E(\theta) = \frac{\alpha}{\alpha + \beta}.$$

$$\hat{\mu}_{\text{MLE}} = \frac{n}{\sum_{i=1}^{n} x_i + n}.$$

Biến đổi một chút ta sẽ có:

$$\begin{aligned}\hat{\theta}_{Bayes} &= \frac{\alpha + n}{\alpha + \beta + \sum_{i=1}^{n} x_i + n} \\ &= \frac{\alpha + \beta}{\alpha + \beta + \sum_{i=1}^{n} x_i} \times \frac{\alpha}{\alpha + \beta} \times \frac{\sum_{i=1}^{n} x_i}{\alpha + \beta + \sum_{i=1}^{n} x_i} \times \frac{n}{\sum_{i=1}^{n} x_i} \\ &= k\ E(\theta) + (1 - k)\hat{\mu}_{\text{MLE}}.\end{aligned}$$

Ta nhận thấy rằng, $\hat{\theta}_{Bayes}$ là tổ hợp lồi của ước lượng hợp lí cực đại và kỳ vọng của phân phối tiên nghiệm.

$$\hat{\theta}_{Bayes} = kE(\theta) + (1 - k)\hat{\mu}_{MLE}.$$

với $k = \frac{\alpha+\beta}{\alpha+\beta+\sum_{i=1}^{n} x_i}$.

- Khi số lượng quan sát càng lớn thì $\hat{\theta}_{Bayes} \approx \hat{\mu}_{MLE}$, tức là lúc này phân phối tiên nghiệm không còn quan trọng nữa.

2.3.2 Khoảng tin cậy Bayes

Định nghĩa 2.3.2 (Khoảng tin cậy Bayes) Một đoạn $[l(d); u(d)]$ dựa trên mẫu $D = d$ đã biết, có độ tin cậy $(1-\alpha)100\%$ của tham số θ nếu:

$$Pr(l(d)) < \theta < u(d)|D = d) = 1 - \alpha$$

Từ định nghĩa trên, ta luôn tìm được khoảng tin cậy Bayes khi có dữ liệu. Có nhiều cách để xác định khoảng tin cậy, như highest posterior density interval (HPD), khoảng tin cậy đối xứng...

3 Ứng dụng

3.1 Số liệu mô phỏng

Để có cái nhìn trực quan hơn về phương pháp Bayes ta sẽ đi vào nghiên cứu mô phỏng dựa trên kết quả của phương pháp suy luận Bayes cho mô hình phân phối hình học, so sánh nó với kết quả của phương pháp ước lượng hợp lý cực đại (MLE).
Sử dụng đầu vào là mô hình dữ liệu phân phối hình học với $\theta = 0.25$ để thực hiện mô phỏng.

Sample size	g1	g2	$\hat{\mu}_{MLE}$	$\hat{\theta}_{Bayes}$
n=10	0.2	0.0004	0.224256	0.202252
	0.2	0.004		0.211904
	0.2	0.04		0.220833
	0.4	0.00004	0.224256	0.387557
	0.4	0.004		0.323076
	0.4	0.04		0.240308
	0.8	0.0004	0.224256	0.741441
	0.8	0.004		0.490476
	0.8	0.04		0.258333

Sample size	g1	g2	$\hat{\mu}_{MLE}$	$\hat{\theta}_{Bayes}$
n=100	0.2	0.0004	0.282485	0.238778
	0.2	0.004		0.274303
	0.2	0.04		0.281792
	0.4	0.00004	0.282485	0.356348
	0.4	0.004		0.299273
	0.4	0.04		0.284122
	0.8	0.0004	0.282485	0.556706
	0.8	0.004		0.333842
	0.8	0.04		0.286834

Sample size	g1	g2	$\hat{\mu}_{MLE}$	$\hat{\theta}_{Bayes}$
n=1000	0.2	0.0004	0.243783	0.239902
	0.2	0.004		0.243371
	0.2	0.04		0.243751
	0.4	0.00004	0.243783	0.263688
	0.4	0.004		0.245998
	0.4	0.04		0.243973
	0.8	0.0004	0.243783	0.293092
	0.8	0.004		0.249022
	0.8	0.04		0.244192

Nhận xét:

Có thể thấy khi ta càng tăng giá trị của g2, giá trị của ước lượng Bayes càng gần với giá trị MLEMLE.

Từ kết quả mô phỏng ta thấy rằng, giá trị ước lượng bằng phương pháp Bayes có xu hướng tiến dần về giá trị ước lượng MLE khi n càng lớn.

3.2 Dữ liệu thật

Ở phần này chúng ta sẽ áp dụng lý thuyết vào giải quyết bài toán với dữ liệu thực tế.
Bộ dữ liệu thực về số lần tung xúc xắc được mặt 6 chấm. Dữ liệu về phân phối hình học được lấy từ thí nghiệm ở ngoài đời thực với cỡ mẫu là 20, tức ta tung xúc xắc đến khi nào được mặt 6 chấm, lặp lại cho đến khi nào đủ 20 lần. Thông tin về bộ dữ liệu như sau:

```
> X
 [1] 1 11 4 2 6 2 0 13 6 5 0 12 0 2 4 0 8 1 1 3
>
```

Bài toán: Giả sử ta muốn tìm tỉ lệ tung được xúc xắc mặt 6 chấm. Ta sẽ đi ước lượng tỉ lệ tung được mặt 6 chấm của xúc xắc.
Chọn tiên nghiệm $\pi(\theta) \sim Beta(\alpha_0, \beta_0)$, kỳ vọng tiên nghiệm bằng 0.166666 kết quả ước lượng như sau:

Sample size	g1	g2	$\hat{\mu}_{MLE}$	$\hat{\theta}_{Bayes}$
n=20	0.166666	0.004	0.198019	0.190171
	0.166666	0.2		0.196277
	0.166666	0.1		0.197899

Ta thấy, khi giá trị g_2 càng trải rộng thì ước lượng Bayes và ước lượng hợp lý cực đại càng gần nhau.

4 Kết luận

Suy luận Bayes cung cấp cho chúng ta một phương pháp tiếp cận mới cho các bài toán ước lượng các tham số của mô hình thống kê. Khác với các phương pháp ước lượng cổ điển là chỉ dựa vào hoàn toàn thông tin từ dữ liệu có được để ước lượng các tham số, thì với phương pháp Bayes ta có thể kết hợp được thêm "niềm tin" từ kinh nghiệm, kiến thức và những suy đoán của mình để lập mô hình ước lượng. Với suy luận Bayes thì việc lựa chọn phân phối tiên nghiệm là một vấn đề quan trọng, nếu chúng ta đưa ra một phân phối tiên nghiệm phù hợp thì kết quả hậu nghiệm thu được sẽ phù hợp, ngược lại, nếu chúng ta có nhiều sai lầm trong việc lựa chọn tiên nghiệm thì kết quả hậu nghiệm có thể không còn tốt.

Tài liệu tham khảo

1. Peter D. Hoff. *"A First Course in Bayesian Statistical Methods"*, University of Washington Department of Statistics (2009).

2. Mary Kathryn Cowles. *"Applied Bayesian Statistics With R and OpenBUGS Examples"*, University of Iowa, Statistics and Actuarial Science (2013).

Phụ lục

CODE R MÔ PHỎNG SỐ LIỆU

```
# Co mau
n<- 10
# Xac suat thanh cong
theta<- 0.25
# Mau
set.seed(0)
x<- rgeom(n, prob = theta)
x
# Uoc luong hop ly cuc dai
s.n<- sum(x); s.n
theta.MLE<- n/(n+s.n); theta.MLE
# Thong tin tien nghiem
g1<- 0.2
g2<- 0.0004
#Tham so tien nghiem
b<- g1-1+g1*(1-g1)^2/g2

a<- b*g1/(1-g1)
# Uoc luong Bayes
theta.Bayes<- (a+n)/(a+b+s.n+n);theta.Bayes
c(theta.MLE, theta.Bayes)
###############
# Co mau
n<- 100
# Xac suat thanh cong
theta<- 0.25
# Mau
set.seed(0)
x<- rgeom(n, prob = theta)
x
# Uoc luong hop ly cuc dai
s.n<- sum(x); s.n
```

```
theta.MLE<- n/(n+s.n); theta.MLE
# Thong tin tien nghiem
g1<- 0.2
g2<- 0.0004
#Tham so tien nghiem
b<- g1-1+g1*(1-g1)^2/g2

<- b*g1/(1-g1)
# Uoc luong Bayes
theta.Bayes<- (a+n)/(a+b+s.n+n);theta.Bayes
c(theta.MLE, theta.Bayes)
##############
# Co mau
n<- 1000
# Xac suat thanh cong
theta<- 0.25
# Mau
set.seed(0)
x<- rgeom(n, prob = theta)
x
# Uoc luong hop ly cuc dai
s.n<- sum(x); s.n
theta.MLE<- n/(n+s.n); theta.MLE
# Thong tin tien nghiem
g1<- 0.2
g2<- 0.0004
#Tham so tien nghiem
b<- g1-1+g1*(1-g1)^2/g2

a<- b*g1/(1-g1)
# Uoc luong Bayes
theta.Bayes<- (a+n)/(a+b+s.n+n);theta.Bayes
c(theta.MLE, theta.Bayes)
##############
```