

Ước lượng tham số sử dụng MLE và MAP

Vũ Lê Thế Anh

TRẠI HÈ PiMA ONLINE 2021



Ngày 31 tháng 7 năm 2021

- 1 Ước lượng tham số
- 2 Ước lượng hợp lý cực đại
- 3 Cực đại hậu nghiệm

Nhắc lại kiến thức

- Một phép thử thì có thể xảy ra 1 trong số nhiều kết quả
- **Sự kiện:** tập hợp một số kết quả có thể
- **Xác suất:**



Nhắc lại kiến thức

- Một phép thử thì có thể xảy ra 1 trong số nhiều kết quả
- **Sự kiện:** tập hợp một số kết quả có thể
- **Xác suất:** gán giá trị $[0, 1]$ cho sự kiện
- **Biến ngẫu nhiên:**



Nhắc lại kiến thức

- Một phép thử thì có thể xảy ra 1 trong số nhiều kết quả
- **Sự kiện:** tập hợp một số kết quả có thể
- **Xác suất:** gán giá trị $[0, 1]$ cho sự kiện
- **Biến ngẫu nhiên:** gán giá trị thực cho kết quả
- **Phân phối:**



Nhắc lại kiến thức

- Một phép thử thì có thể xảy ra 1 trong số nhiều kết quả
- **Sự kiện**: tập hợp một số kết quả có thể
- **Xác suất**: gán giá trị $[0, 1]$ cho sự kiện
- **Biến ngẫu nhiên**: gán giá trị thực cho kết quả
- **Phân phối**: một cách phân chia xác suất cho tất cả sự kiện, đặc trưng cho một biến ngẫu nhiên



Table of Contents

- 1 Ước lượng tham số
- 2 Ước lượng hợp lý cực đại
- 3 Cực đại hậu nghiệm

Một số câu hỏi (có thể) đã gặp

Tung đồng xu #1

Một đồng xu có xác suất ra mặt ngửa là 80%, hỏi xác suất ra mặt ngửa khi tung đồng xu này?



Một số câu hỏi (có thể) đã gặp

Tung đồng xu #1

Một đồng xu có xác suất ra mặt ngửa là 80%, hỏi xác suất ra mặt ngửa khi tung đồng xu này?

Tung đồng xu #2

Thực hiện phép thử tung một đồng xu. Gọi X là biến ngẫu nhiên nhận giá trị 1 khi kết quả là mặt ngửa và 0 trong các trường hợp khác. Giả sử X tuân theo phân phối Bernoulli với tham số $p = 0.8$.

Yêu cầu: Tính $P(X = 1)$.



Một số câu hỏi (có thể) đã gặp

Tung đồng xu #1

Một đồng xu có xác suất ra mặt ngửa là 80%, hỏi xác suất ra mặt ngửa khi tung đồng xu này?

Tung đồng xu #2

Thực hiện phép thử tung một đồng xu. Gọi X là biến ngẫu nhiên nhận giá trị 1 khi kết quả là mặt ngửa và 0 trong các trường hợp khác. Giả sử X tuân theo phân phối Bernoulli với tham số $p = 0.8$.

Yêu cầu: Tính $P(X = 1)$.

Trả lời: $X \sim \text{Bernoulli}(p)$ với $p = 0.8$,

$$f_p(x) = \begin{cases} p, & x = 1 \\ 1 - p, & x = 0 \\ 0, & x \text{ khác} \end{cases}, \quad P(X = 1) = f_{0.8}(1)$$



Một số câu hỏi (có thể) đã gặp

Do chiều cao

Chọn ngẫu nhiên một người Việt Nam và đo chiều cao (mét) của người đó. Gọi X là biến ngẫu nhiên có giá trị bằng chiều cao đo được. Giả sử X tuân theo phân phối chuẩn với tham số $\mu = 1.65$ và $\sigma^2 = 0.3$.

Yêu cầu: Tính xác suất người này thấp hơn 1 mét, $P(0 \leq X < 1)$.



Một số câu hỏi (có thể) đã gặp

Đo chiều cao

Chọn ngẫu nhiên một người Việt Nam và đo chiều cao (mét) của người đó. Gọi X là biến ngẫu nhiên có giá trị bằng chiều cao đo được. Giả sử X tuân theo phân phối chuẩn với tham số $\mu = 1.65$ và $\sigma^2 = 0.3$.

Yêu cầu: Tính xác suất người này thấp hơn 1 mét, $P(0 \leq X < 1)$.

Trả lời: $X \sim \mathcal{N}(\mu, \sigma^2)$ với $\mu = 1.65$ và $\sigma^2 = 0.3$

$$f_{\mu, \sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad P(X < 1) = \int_0^1 f_{\mu, \sigma^2}(x) dx$$



Điểm chung

Xét biến ngẫu nhiên X tuân theo phân phối $D(\theta)$, tức biết:

- Họ (family) của phân phối
- Các tham số θ

Chúng ta muốn:

- Truy vấn về kết quả (xác suất, kỳ vọng,...)



Một số câu hỏi... khác #1

Tung đồng xu #1

Trung nhật được một đồng xu trên đường. Anh tung thử 10 lần thì thu được kết quả lần lượt là $H, T, H, H, H, H, T, T, T, H$. Hỏi xác suất tung được mặt ngửa khi tung đồng xu trên?



Một số câu hỏi... khác #1

Tung đồng xu #1

Trung nhật được một đồng xu trên đường. Anh tung thử 10 lần thì thu được kết quả lần lượt là $H, T, H, H, H, H, T, T, T, H$. Hỏi xác suất tung được mặt ngửa khi tung đồng xu trên?

Tung đồng xu #2

Thực hiện phép thử tung một đồng xu. Gọi X là biến ngẫu nhiên có giá trị 1 khi kết quả là mặt ngửa và 0 trong các trường hợp khác. Giả sử X tuân theo phân phối D .

Lấy 10 mẫu ngẫu nhiên từ D ta thu được $\{1, 0, 1, 1, 1, 1, 0, 0, 0, 1\}$.

Yêu cầu: tìm phân phối D .



Lấy mẫu ngẫu nhiên

- Phân phối D xác định xác suất cho mỗi giá trị (hoặc khoảng giá trị)



Lấy mẫu ngẫu nhiên

- Phân phối D xác định xác suất cho mỗi giá trị (hoặc khoảng giá trị)
- Lấy mẫu từ phân phối: sinh ra giá trị phản ánh phân phối

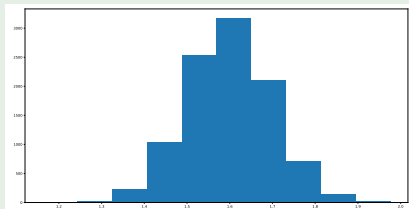


Một số câu hỏi... khác #2

Đo chiều cao

Mời ngẫu nhiên 3000 người đến đo chiều cao. Gọi X là biến ngẫu nhiên có giá trị bằng chiều cao đo được. Giả sử X tuân theo phân phối D .

Kết quả thu được cho biểu đồ histogram như sau



Yêu cầu: tìm phân phối D .



Điểm chung

Xét một phân phối D , biết:

- Các mẫu lấy từ D

Muốn: tìm D



Điểm chung

Xét một phân phối D , biết:

- Các mẫu lấy từ D

Muốn: tìm $D \Rightarrow$ **khó!**



Điểm chung

Xét một phân phối D , biết:

- Các mẫu lấy từ D

Muốn: tìm $D \Rightarrow$ **khó!**

Giả sử biết thêm **họ** của phân phối, $D(\theta)$, ta cần tìm:

- Các tham số θ



Bài toán ước lượng tham số

Bài toán ước lượng tham số

Cho một phân phối xác suất $D(\theta)$ được tham số hóa bởi θ chưa biết. Cho x_1, x_2, \dots, x_n là n giá trị thu thập bằng cách lấy mẫu từ $D(\theta)$. Xác định θ .



Bài toán ước lượng tham số

Bài toán ước lượng tham số

Cho một phân phối xác suất $D(\theta)$ được tham số hóa bởi θ chưa biết. Cho x_1, x_2, \dots, x_n là n giá trị thu thập bằng cách lấy mẫu từ $D(\theta)$. Xác định θ .

Hai hướng tiếp cận nổi bật:

- Hợp lý cực đại (Maximum Likelihood Estimation, MLE)
- Cực đại hậu nghiệm (Maximum A Posteriori, MAP)



Ngoài lề

Câu hỏi thường hỏi

Làm thế nào để biết họ phân phối?



Table of Contents

- 1 Ước lượng tham số
- 2 Ước lượng hợp lý cực đại
- 3 Cực đại hậu nghiệm



Hợp lý (Likelihood)

Hợp lý

Gọi X là biến ngẫu nhiên tuân theo phân phối $D(\theta)$. Cho x là một giá trị cụ thể, likelihood của x được định nghĩa là một hàm theo θ , ký hiệu $L(\theta|x)$, thỏa:

$$L(\theta|x) = f_{\theta}(x)$$



Ví dụ

Hợp lý với phân phối Poisson

Cho biến ngẫu nhiên $X \sim \text{Pois}(2.5)$. Tìm hợp lý trong trường hợp $x = 0, 1, \dots, 6$.

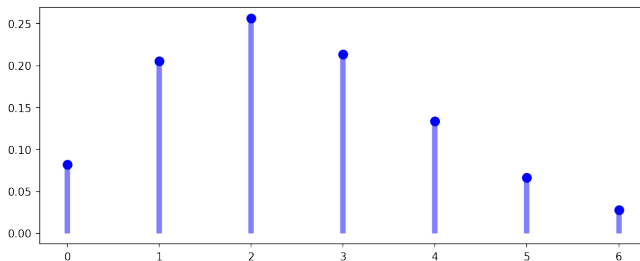


Ví dụ

Hợp lý với phân phối Poisson

Cho biến ngẫu nhiên $X \sim \text{Pois}(2.5)$. Tìm hợp lý trong trường hợp $x = 0, 1, \dots, 6$.

Giải



Ví dụ (khác)

Hợp lý với phân phối Poisson

Cho biến ngẫu nhiên $X \sim \text{Pois}(\lambda)$ và giá trị cụ thể $x = 3$. Tính giá trị hợp lý tại $\lambda = 0, 2.5, 5$.

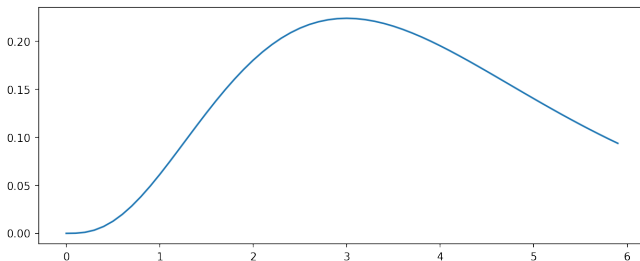


Ví dụ (khác)

Hợp lý với phân phối Poisson

Cho biến ngẫu nhiên $X \sim \text{Pois}(\lambda)$ và giá trị cụ thể $x = 3$. Tính giá trị hợp lý tại $\lambda = 0, 2.5, 5$.

Giải



Hợp lý của nhiều giá trị

Giả sử: x_1, \dots, x_n là n giá trị của X lấy mẫu **độc lập** và **cùng từ** $D_X(\theta)$.

$$L(\theta|x_1, x_2, \dots x_n) = f_{\theta}(x_1)f_{\theta}(x_2) \dots f_{\theta}(x_n) = \prod_{i=1}^n f_{\theta}(x_i)$$



Ước lượng hợp lý cực đại

Ý tưởng: tham số “hợp lý” nhất với dữ liệu đang có



Ước lượng hợp lý cực đại

Ý tưởng: tham số “hợp lý” nhất với dữ liệu đang có

Ước lượng hợp lý cực đại

Có:

- $D(\theta)$: phân phối tham số hóa bởi θ
- x_1, \dots, x_n : các giá trị được giả sử là lấy mẫu từ $D(\theta)$

Tham số tìm được theo phương pháp MLE:

$$\theta_{MLE} = \operatorname{argmax}_{\theta} L(\theta | x_1, \dots, x_n)$$



Ước lượng hợp lý cực đại

Ý tưởng: tham số “hợp lý” nhất với dữ liệu đang có

Ước lượng hợp lý cực đại

Có:

- $D(\theta)$: phân phối tham số hóa bởi θ
- x_1, \dots, x_n : các giá trị được giả sử là lấy mẫu từ $D(\theta)$

Tham số tìm được theo phương pháp MLE:

$$\theta_{MLE} = \operatorname{argmax}_{\theta} L(\theta | x_1, \dots, x_n)$$

Trong nhiều trường hợp, ta sử dụng hàm **đồng biến** log,

$$\theta_{MLE} = \operatorname{argmax}_{\theta} \log L(\theta | x_1, \dots, x_n)$$



Ví dụ

Tung đồng xu

Trung nhật được một đồng xu trên đường. Anh tung thử 10 lần thì thu được kết quả lần lượt là $H, T, H, H, H, H, T, T, T, H$. Hỏi xác suất tung được mặt ngửa khi tung đồng xu trên?



Ví dụ

Tung đồng xu

Trung nhật được một đồng xu trên đường. Anh tung thử 10 lần thì thu được kết quả lần lượt là $H, T, H, H, H, H, T, T, T, H$. Hỏi xác suất tung được mặt ngửa khi tung đồng xu trên?

- Chọn $D_X(\theta)$ là phân phối **Bernoulli** tham số p

$$f_p(x) = \begin{cases} p, & x = 1 \\ 1 - p, & x = 0 \\ 0, & x \text{ khác} \end{cases}$$

- Hàm likelihood

$$L(p|1, 0, 1, 1, 1, 1, 0, 0, 0, 1) =$$



Ví dụ

Tung đồng xu

Trung nhật được một đồng xu trên đường. Anh tung thử 10 lần thì thu được kết quả lần lượt là $H, T, H, H, H, H, T, T, T, H$. Hỏi xác suất tung được mặt ngửa khi tung đồng xu trên?

- Chọn $D_X(\theta)$ là phân phối **Bernoulli** tham số p

$$f_p(x) = \begin{cases} p, & x = 1 \\ 1 - p, & x = 0 \\ 0, & x \text{ khác} \end{cases}$$

- Hàm likelihood

$$L(p|1, 0, 1, 1, 1, 1, 0, 0, 0, 1) = p^6(1 - p)^4$$



Ví dụ

■ Hàm log-likelihood

$$\log L(p|1, 0, 1, 1, 1, 1, 0, 0, 0, 1) = 6 \log p + 4 \log(1 - p)$$



Ví dụ

- Hàm log-likelihood

$$\log L(p|1, 0, 1, 1, 1, 1, 0, 0, 0, 1) = 6 \log p + 4 \log(1 - p)$$

- Ước lượng hợp lý cực đại

$$p_{MLE} = \operatorname{argmax}_p \log L(p|1, 0, 1, 1, 1, 1, 0, 0, 0, 1)$$



Ví dụ

- Hàm log-likelihood

$$\log L(p|1, 0, 1, 1, 1, 1, 0, 0, 0, 1) = 6 \log p + 4 \log(1 - p)$$

- Ước lượng hợp lý cực đại

$$p_{MLE} = \underset{p}{\operatorname{argmax}} \log L(p|1, 0, 1, 1, 1, 1, 0, 0, 0, 1)$$

- Kết quả

$$p_{MLE} = \frac{6}{6 + 4} = 0.6$$



Bài tập

- 1 Trên chiến trường thể chiến 2, phe Đồng Minh tìm thấy 4 chiếc xe tăng Đức đánh số hiệu lần lượt là 19, 40, 42, 60. Ước đoán số lượng xe tăng Đức theo tinh thần của MLE.
- 2 Giải trường hợp tổng quát của ví dụ “Tung đồng xu” trên: Cho phân phối Bernoulli với tham số p và n quan sát x_1, x_2, \dots, x_n . Ước lượng p theo phương pháp MLE.
- 3 Họ phân phối chuẩn có 2 tham số là μ và σ^2 . Giả sử có n quan sát x_1, x_2, \dots, x_n . Ước lượng μ và σ^2 theo phương pháp MLE.



Table of Contents

- 1 Ước lượng tham số
- 2 Ước lượng hợp lý cực đại
- 3 Cực đại hậu nghiệm



Động lực

Thiếu sót của MLE:

- Xem tham số như một giá trị cố định
- Không thể thêm tri thức bên ngoài vào



Động lực

Thiếu sót của MLE:

- Xem tham số như một giá trị cố định
- Không thể thêm tri thức bên ngoài vào

(lại là) Tung đồng xu

Tình huống cũ: Trung nhặt được một đồng xu trên đường. Anh tung thử 10 lần thì thu được kết quả lần lượt là $H, T, H, H, H, H, T, T, T, H$.

Yếu tố mới: Thế Anh, anh họ Trung, nhìn thấy đồng xu quen thuộc mới bảo Trung rằng đó là đồng xu anh đã vứt đi. Thế Anh bảo anh dùng nó để đi lừa người khác, nên cố tình làm cho mặt sấp dễ ra hơn mặt ngửa.



Ý tưởng

- Xem **tham số** là một **biến ngẫu nhiên** (có phân phối)
- Dùng **tri thức** và **quan sát** để “cập nhật”



Định lý Bayes

$$P(A|B) =$$



Định lý Bayes

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \propto P(B|A)P(A)$$

- $P(A)$: xác suất tiên nghiệm (prior)
- $P(B|A)$: xác suất hậu nghiệm (posterior)



Định lý Bayes cho biến ngẫu nhiên

$$f_{\Theta|x}(\theta) = \frac{f_{X|\theta}(x)f_{\Theta}(\theta)}{f_X(x)} \propto f_{X|\theta}(x)f_{\Theta}(\theta)$$



Định lý Bayes cho biến ngẫu nhiên

$$f_{\Theta|x}(\theta) = \frac{f_{X|\theta}(x)f_{\Theta}(\theta)}{f_X(x)} \propto f_{X|\theta}(x)f_{\Theta}(\theta)$$

- f_{Θ} : hàm khối/mật độ của **phân phối tiên nghiệm** (prior)
- $f_{X|\theta}$: hàm hợp lý (likelihood), $f_{X|\theta}(x) = L(\theta|x)$
- $f_{\Theta|x}$: hàm khối/mật độ của **phân phối hậu nghiệm** (posterior)



Ví dụ

Xác suất hậu nghiệm với phân phối Poisson

Cho biến ngẫu nhiên $X \sim \text{Pois}(\lambda)$ và giá trị cụ thể $x = 3$. Cho λ là một giá trị cụ thể của $\Lambda \sim \text{Exp}(2)$. Tìm hàm mật độ hậu nghiệm.

$$\Lambda \sim \text{Exp}(2) \Rightarrow f_{\Lambda}(\lambda) = 2e^{-2\lambda}$$

Giải

$$f_{\Lambda|X=3}(\lambda) \propto$$



Ví dụ

Xác suất hậu nghiệm với phân phối Poisson

Cho biến ngẫu nhiên $X \sim \text{Pois}(\lambda)$ và giá trị cụ thể $x = 3$. Cho λ là một giá trị cụ thể của $\Lambda \sim \text{Exp}(2)$. Tìm hàm mật độ hậu nghiệm.

$$\Lambda \sim \text{Exp}(2) \Rightarrow f_{\Lambda}(\lambda) = 2e^{-2\lambda}$$

Giải

$$f_{\Lambda|X=3}(\lambda) \propto f_{X|\Lambda}(\lambda)(3)f_{\Lambda}(\lambda) = \left(e^{-\lambda} \frac{\lambda^3}{3!} \right) (2e^{-2\lambda})$$

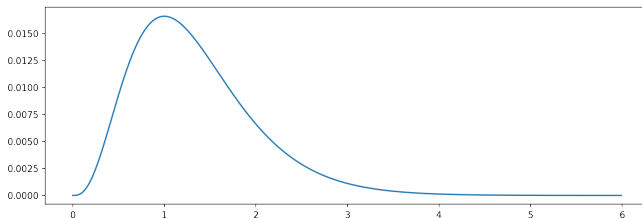


Ví dụ

Xác suất hậu nghiệm với phân phối Poisson

Cho biến ngẫu nhiên $X \sim \text{Pois}(\lambda)$ và giá trị cụ thể $x = 3$. Cho λ là một giá trị cụ thể của $\Lambda \sim \text{Exp}(2)$. Tìm hàm mật độ hậu nghiệm.

Giải



Cực đại hậu nghiệm

Cực đại hậu nghiệm (Maximum A Posteriori)

Có:

- $D(\theta)$: phân phối tham số hóa bởi θ
- x_1, \dots, x_n : các giá trị được giả sử là lấy mẫu từ $D(\theta)$
- $D_{\Theta}(\omega)$: phân phối của tham số, tham số hóa bởi ω

Tham số tìm được theo phương pháp MAP:

$$\theta_{MAP} = \underset{\theta}{\operatorname{argmax}} L(\theta|x_1, \dots, x_n) f_{\Theta}(\theta)$$



Ví dụ

Tung đồng xu

Trung nhật được một đồng xu trên đường và tung thử 10 lần thì thu được chuỗi $H, T, H, H, H, H, T, T, T, H$. Hỏi xác suất tung được mặt ngửa của đồng xu, biết nó không quá 0.5?

- Đề xuất phân bố tiên nghiệm: $P \sim U(0, 0.5)$

$$f_P(p) = \begin{cases} 2, & p \in [0, 0.5] \\ 0, & p \text{ khác} \end{cases}$$

- Hàm hậu nghiệm tỉ lệ với

$$L(p|1, 0, 1, 1, 1, 1, 0, 0, 0, 1)f_P(p) = \begin{cases} 2p^6(1-p)^4, & p \in [0, 0.5] \\ 0, & p \text{ khác} \end{cases}$$



Ví dụ

Tung đồng xu

Trung nhật được một đồng xu trên đường và tung thử 10 lần thì thu được chuỗi $H, T, H, H, H, H, T, T, T, H$. Hỏi xác suất tung được mặt ngửa của đồng xu, biết nó không quá 0.5?

- Cực đại hóa xác suất hậu nghiệm \equiv tìm $p \in [0, 0.5]$ để tối ưu

$$2p^6(1-p)^4 \text{ hoặc } 6 \log p + 4 \log(1-p) + \text{log 2}$$



Ví dụ

Tung đồng xu

Trung nhật được một đồng xu trên đường và tung thử 10 lần thì thu được chuỗi $H, T, H, H, H, H, T, T, T, H$. Hỏi xác suất tung được mặt ngửa của đồng xu, biết nó không quá 0.5?

- Cực đại hóa xác suất hậu nghiệm \equiv tìm $p \in [0, 0.5]$ để tối ưu

$$2p^6(1-p)^4 \text{ hoặc } 6 \log p + 4 \log(1-p) + \text{log 2}$$

- Kết quả: $p_{MAP} = 0.5 \neq p_{MLE}$.



Bài tập

- 1 Một họ phân phối thường được chọn làm phân phối tiên nghiệm cho các tham số thuộc loại xác suất (như p trong phân phối Bernoulli) là phân phối Beta.

$$f_{a,b}(p) = \frac{1}{B(a,b)} p^{a-1} (1-p)^{b-1}, \quad B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

Tìm công thức p_{MAP} trong trường hợp sử dụng phân phối tiên nghiệm trên (giải tổng quát với a, b là hai tham số tự đặt).



Note

- Chúng ta vừa: **cập nhật** lại phân phối của tham số **từ phân phối tiên nghiệm** khi **có thêm dữ liệu**
- MAP: chọn ra **giá trị có khả năng nhất** (hàm mật độ tại đó cao nhất) trong phân phối này
- Ngoài ra: kỳ vọng, phương sai, suy luận nói chung về tham số

