Ước lượng Bayes (tiếp):

1. Ước lượng điểm (tiếp):

- Sau khi tìm hiểu về ước lượng hàm Bayes quardatic Loss, ta sẽ tìm hiểu về một số loại ước lượng điểm khác.
- Có các dạng ước lượng khác như:

❖Ước lượng tuyệt đối.

❖Ước lượng 0-1.

a) Ước lượng tuyệt đối:

- Ta sẽ tìm hiểu về ước lượng tuyệt đối.
- Xét hàm loss:

$$L\left(\hat{\theta}, \theta\right) = |\hat{\theta} - \theta|$$

- Ở đây theta là tham số của bài toán và theta mu là điểm cần xét.

- Như vậy ta có làm loss kì vọng là:

$$\rho\left(\hat{\theta},\theta\right) = \mathrm{E}\left(\mathrm{L}\left(\hat{\theta},\theta\right)\right) = \int \left|\hat{\theta} - \theta\right| \mathrm{p}(\theta|\mathrm{D})\mathrm{d}\theta$$

Chú ý rằng ở dạng tổng quát thì θ và $\hat{\theta}$ là các vector ngẫu nhiên.

- Ta sẽ tìm hiểu bài toán trong trường hợp θ và $\hat{\theta}$ là các số thực dương.
- Trước tiên ta lấy khoảng (m, n) cho giá trị θ:

$$\rho(\hat{\theta}, \theta) = \int_{\mathrm{m}}^{\mathrm{n}} \left|\hat{\theta} - \theta\right| \mathrm{p}(\theta|\mathrm{D}) \mathrm{d}\theta$$

- Trong đó $\mathrm{m}, \mathrm{n} > 0$.

- Như vậy, ta sẽ có phân tích:

$$\int_{m}^{n} \left|\theta - \hat{\theta}\right| \mathrm{p}(\theta|\mathrm{D}) \mathrm{d}\theta$$

$$= \int_{\hat{\theta}}^{n} \left(\theta - \hat{\theta}\right) \mathrm{p}(\theta|\mathrm{D}) \mathrm{d}\theta + \int_{\hat{\theta}}^{m} (\hat{\theta} - \theta) \mathrm{p}(\theta|\mathrm{D}) \mathrm{d}\theta$$

- Biểu thức trên bằng với:

$$\mathrm{H}(\theta) = \hat{\theta}\mathrm{F}\left(\hat{\theta}\right) - \int_{\hat{\theta}}^{\mathrm{n}} \theta . p(\theta|\mathrm{D}) d\theta + \int_{\hat{\theta}}^{\mathrm{m}} \theta . p(\theta|a) d\theta + \hat{\theta}\left(1 - \mathrm{F}\left(\hat{\theta}\right)\right)$$

Trong đó $\int_{\mathrm{m}}^{\mathrm{n}} \mathrm{p}(\theta|\mathrm{D})\mathrm{d}\theta = \mathrm{F}(\theta)$

- Ở đây, ta cần tìm cực trị của hàm số trên để xác định xem với θ như thế nào thì hàm Loss sẽ đạt min.
- Vì đây là hàm một biến nên ta hoàn toàn có thể áp dụng các kiến thức tìm cực trị hàm số cơ bản.
- Chú ý rằng $\mathrm{F}(\theta)$ là một đại lượng độc lập vì đã được tính toán dựa trên các khoảng bị chặn nên ta không phải để ý đến nó.

Bước tìm min.

- Tính $\mathrm{H}'\left(\hat{\theta}\right)$:

Với $\mathrm{H}(\hat{\theta}) = \hat{\theta}\mathrm{F}\left(\hat{\theta}\right) - \int_{\hat{\theta}}^{\mathrm{n}} \theta.(\theta|\mathrm{D})d\theta + \int_{\hat{\theta}}^{\mathrm{m}} \theta.p(\theta|\mathrm{D})d\theta + \hat{\theta}\left(1 - \mathrm{F}\left(\hat{\theta}\right)\right)$, ta có:

$$\mathrm{H}'(\theta) = \mathrm{F}\left(\hat{\theta}\right) + \hat{\theta}\mathrm{F}'\left(\hat{\theta}\right) + 1 - \mathrm{F}(\theta) - \hat{\theta}\mathrm{F}'\left(\hat{\theta}\right)$$

$$= 1 - 2\mathrm{F}(\hat{\theta})$$

- Như vậy để hàm $\mathrm{H}(\hat{\theta})$ đạt cực trị, ta cần có:

$$\mathrm{H}'\left(\hat{\theta}\right) = 0$$

- Nói cách khác, ta cần $\mathrm{F}(\theta) = \frac{1}{2}$.
- Hàm này đạt cực tiểu tại một gia trị $\hat{\theta}$ sao cho $\mathrm{F}\left(\hat{\theta}\right) = \frac{1}{2}$.

- Lưu ý:

Ta có hàm:

$$\mathrm{H}(\theta) = \hat{\theta}\mathrm{F}\big(\hat{\theta}\big) - \int_{\widehat{\theta}}^{\mathrm{n}} \theta.\, p(\theta|\mathrm{D})d\theta + \int_{\widehat{\theta}}^{\mathrm{m}} \theta.\, p(\theta|a)d\theta + \hat{\theta}\big(1 - \mathrm{F}(\theta)\big)$$

$$= \hat{\theta}\mathrm{F}\big(\hat{\theta}\big) + \hat{\theta}\left(1 - \mathrm{F}\big(\hat{\theta}\big)\right) + \mathrm{a} - \mathrm{b}$$

Là một hàm lồi.

- Như vậy ta thấy $\mathrm{H}(\theta)$ là hàm lồi và đạt cực tiểu tại $\hat{\theta}$ sao cho $\mathrm{F}\left(\hat{\theta}\right) = \frac{1}{2}$.
- Giá trị $\hat{\theta}$ sao cho hàm mật độ xác suất $\mathrm{F}\left(\hat{\theta}\right) = \frac{1}{2}$ còn được gọi là giá trị trung vị - "median".
- Như vậy trong trường hợp sử dụng hàm loss tuyệt đối, nó sẽ đạt cực tiểu nếu θ là median của phân phối hậu nghiệm $\theta|\mathrm{D}$.

Định lý:

- Ta có một định lý tổng quát hơn:

  Định lý: Cho mẫu ngẫu nhiên $\mathrm{X}_1, \mathrm{X}_2, \dots, \mathrm{X}_\mathrm{n}$ với hàm mật độ xác suất $\mathrm{f}(\mathrm{X}|\theta)$ trong đó θ là tham số cần ước lượng. Nếu ta sử dụng hàm loss là hàm loss tuyệt đối, thì ước lượng Bayes $\hat{\theta}$ chính là median (trung vị) của phân phối xác suất hậu nghiệm.

Ví dụ minh hoạ:

- Cho biến ngẫu nhiên X có phân phối nhị thức với $\mathrm{n} = 3$ và tham số θ. Ta có phân phối xác suất của bộ dữ liệu này là:

$$\mathrm{p}(\mathrm{X}|\theta) = \mathrm{C}_3^{\mathrm{x}}\theta^{\mathrm{x}}(1-\theta)^{3-\mathrm{x}}$$

- Bước 1: chọn phân phối xác suất tiên nghiệm cho theta. Giả sử ta có phân phối đều:

$$\begin{cases} \mathrm{p}(\theta) = 1; \theta \in (½, 1) \\ \mathrm{p}(\theta) = 0; \theta \notin (\frac{1}{2}, 1) \end{cases}$$

- Khi đó tại x = 3, ta có tử của phân phối xác suất hậu nghiệm:

$$\mathrm{P}(\mathrm{X}, \theta) = p(\theta).\, \mathrm{p}(\mathrm{X}_3|\theta) = \theta^3; \; \frac{1}{2} < \theta < 1$$

- Lưu ý ta hoàn toàn có thể tính toán theo từng biến ngẫu nhiên trong quá trình trên.

- Bước 2: tính $\mathrm{F}(\theta|\mathrm{D})$ (chính là $\mathrm{F}(\theta|\mathrm{X}_3)$).

$$\mathrm{P}(\mathrm{X}_3) = \int_{1/2}^{1} \mathrm{p}(\theta, X)\mathrm{d}\theta = \int_{1/2}^{1} \theta^3 \mathrm{d}\theta = ?$$

Hãy tính con số theta.

- Ta tính được: $\mathrm{F}(\mathrm{X}_3) = \frac{15}{64}$
- Như vậy ta có hàm mật độ:

$$\begin{cases} \mathrm{F}(\mathrm{X}_3) = \dfrac{15}{64}, \forall\theta \in (\dfrac{1}{2}, 1) \\ 0 \end{cases}$$

- Như vậy ta có:

$$\mathrm{P}(\theta|\mathrm{x}=3)=\frac{P(X_3,\theta)}{\mathrm{F}(\mathrm{X}_3)}=\frac{15}{64}\theta^3, ½<\theta<1$$

$$\mathrm{P}(\theta|\mathrm{x}=3)=0, \theta\notin(\frac{1}{2},1)$$

- Bước 3: tính ước lượng Bayes.
- Theo định lý ở trên, ta có

$$\hat{\theta} = \mathrm{med}\big(P(\theta|\mathrm{x} = 3)\big)$$

- Ta có:

$$\frac{1}{2} = \int_{\frac{1}{2}}^{\widehat{\theta}} \frac{64}{15}\theta^3 \mathrm{d}\theta = \frac{64}{15}(\hat{\theta}^4 - \frac{1}{16})$$

- Ta tính được ước lượng Bayes:

$$\hat{\theta} = 0{,}8537$$

- Như vậy ta tìm được ước lượng cho trường hợp $\mathrm{X} = 3$.
- Bài tập: tìm ước lượng trong trường hợp tổng quát (tính với X=0, X=1, X=2)

b) Ước lượng 0 – 1:

- Xét hàm loss:

$$L(\hat{\theta},\theta) = 0,\ \ \text{nếu}\ \left|\hat{\theta}-\theta\right| < \text{a};\ \ L(\hat{\theta},\theta) = 1,\ \text{nếu}\ \left|\hat{\theta}-\theta\right| \geq \text{a}$$

- Ở đây theta là tham số của bài toán và theta mũ là điểm cần xét.

- Ta sẽ được hàm loss kì vọng:

$\rho(\hat{\theta}, \theta) = \mathrm{E}\left(\mathrm{L}(\hat{\theta}, \theta)\right) = \int \quad \mathrm{p}(\theta|\mathrm{D})\mathrm{d}\theta$ nếu $\left|\hat{\theta} - \theta\right| \geq \mathrm{a}$,

$\rho(\hat{\theta}, \theta) = 0$ nếu $\left|\hat{\theta} - \theta\right| < \mathrm{a}$.

- Ta không để ý tới trường hợp bằng 0.

- Ta sẽ tìm hiểu bài toán trong trường hợp $\hat{\theta}, \theta$ và a là các số thực dương.
- Trước tiên ta lấy khoảng (m, n) cho giá trị θ:

$$\rho\left(\hat{\theta}, \theta\right) = \int_{\mathrm{m}}^{\mathrm{n}} a.\,\mathrm{p}(\theta|\mathrm{D})\mathrm{d}\theta =$$

- Trong đó $\mathrm{m} > 0$.

- Phân tách biểu thức trên, ta có:

$$\rho\left(\hat{\theta}, \theta\right) = \int_{\mathrm{m}}^{\mathrm{n}} \mathrm{p}(\theta|\mathrm{D})\mathrm{d}\theta$$

$$= \int_{\mathrm{m}}^{\widehat{\theta}-\mathrm{a}} \mathrm{p}(\theta|\mathrm{D})\mathrm{d}\theta + \int_{\mathrm{n}}^{\widehat{\theta}+\mathrm{a}} \mathrm{p}(\theta|\mathrm{D})\mathrm{d}\theta$$

- Biểu thức vừa rồi sẽ bằng:

$$\rho\left(\hat{\theta}, \theta\right) = 1 - \int_{\widehat{\theta}-\mathrm{a}}^{\widehat{\theta}+\mathrm{a}} \mathrm{p}(\theta|\mathrm{D})\mathrm{d}\theta$$

- Bài toán ước lượng điểm yêu cầu ta tìm cực tiểu của $\rho\left(\hat{\theta}, \theta\right)$, cũng chính là tìm cực đại của $\int_{\widehat{\theta}-\mathrm{a}}^{\widehat{\theta}+\mathrm{a}} \mathrm{p}(\theta|\mathrm{D})\mathrm{d}\theta$.

- Làm thế nào để tìm cực đại của hàm :

$$\mathrm{F}\left(\hat{\theta},\theta\right)=\int_{\widehat{\theta}-\mathrm{a}}^{\widehat{\theta}+\mathrm{a}}\mathrm{p}(\theta|\mathrm{D})\mathrm{d}\theta \text{ ?}$$

- Chú ý rằng ta sẽ tìm $\hat{\theta}$ sao cho $\mathrm{F}(\hat{\theta},\theta)$ đạt max.

- Khi $a \to 0$, ta được :

$$F(\hat{\theta}, \theta) = F(\hat{\theta} + a \middle| D) - F(\hat{\theta} - a \middle| D) \to 0$$

Biểu thức trên tương đương với:

$$\frac{F(\hat{\theta}, \theta)}{2a} = \frac{F(\hat{\theta} + a|D) - F(\hat{\theta} - a|D)}{\hat{\theta} + a - (\hat{\theta} - a)} \sim F'(\hat{\theta}|D) = p(\hat{\theta}, D)$$

- Mà ta có:

$$\hat{\theta} = \underset{\hat{\theta}}{\mathrm{Argmax}}\,\mathrm{F}\left(\hat{\theta}, \theta\right)$$

$$= \mathrm{Argmax}_{\hat{\theta}} \frac{\mathrm{F}\left(\hat{\theta}, \theta\right)}{2\mathrm{a}}$$

$$= \mathrm{Argmax}_{\hat{\theta}}\ \mathrm{p}(\hat{\theta}|\mathrm{D})$$

- Như vậy, trong trường hợp sử dụng hàm loss 0-1 để ước lượng, bài toán ước lượng Bayes lại trở thành bài toán tìm MAP.
- Bài toán tìm MAP có thể xem lại ở các slide đầu tiên.
- Như vậy ta có định lý:
  Định lý: Cho mẫu ngẫu nhiên $\mathrm{X}_1, \mathrm{X}_2, \dots, \mathrm{X}_\mathrm{n}$ với hàm mật độ xác suất $\mathrm{p}(\mathrm{X}|\theta)$ trong đó θ là tham số cần ước lượng. Nếu ta sử dụng hàm loss là hàm loss 0-1, thì ước lượng Bayes $\hat{\theta}$ chính là MAP của phân phối xác suất hậu nghiệm.

Ví dụ:

- Ta sẽ tìm hiểu với bài toán với phân phối bernouli (mức độ đơn giản nhất).
- Một học viên cao học nộp đơn tìm vị trí tiến sĩ sau khi tốt nghiệp. Biết rằng trước anh thì rất nhiều người đã nộp đơn ở 5 nơi, và đa số chỉ có 1 lần thành công và 4 lần thất bại. Hiện tại anh vẫn đang tiếp tục thử ở những nơi khác để tìm ra trường Đại Học phù hợp nhất. Anh ta nộp ở 3 nơi, và bản thân anh ấy dự đoán rằng mình chỉ có thể có 1 lần thành công. Hãy sử dụng hàm loss 0-1 để ước lượng khả năng thành công 1 trong 3 lần thử sắp tới của anh ấy.

- Gọi θ là tham số “xác suất thành công” của thạc sĩ trên.
- Bài toán theo phân phối Bernoulli, nhưng khi sử dụng phương pháp Bayes thì ta sẽ cho phân phối xác suất Beta cho tham số.
- Khi đó: $\theta \sim \beta(1,4)$. Theo công thức của phân phối beta, ta có:

$$\mathrm{P}(\theta) = \frac{\Gamma(\alpha_1 + \alpha_0)}{\Gamma(\alpha_1)\Gamma(\alpha_0)} . \theta^{\alpha_1 - 1}(1 - \theta)^{\alpha_0} = 4 . \theta^0 . (1 - \theta)^3$$

- Theo mô tả của bài toán, ta có $\mathrm{D}|\theta \sim \mathrm{B}(1,2)$:

$$\mathrm{P}(\mathrm{D}|\theta) = \mathrm{C}_3^1.\theta.(1-\theta)^2$$

- Như vậy, ta có xác suất hậu nghiệm:

$$\mathrm{P}(\theta|\mathrm{D}) \sim \mathrm{P}(\theta).\mathrm{P}(\mathrm{D}|\theta) = 12.\theta(1-\theta)^5$$

- Vì sử dụng hàm loss 0-1, theo định lý vừa có, ta được:

$$\hat{\theta} = \mathrm{Argmax}_{\theta}\theta(1-\theta)^5$$

- Đây chính là bài toán tìm cực trị của $\theta(1-\theta)^5$.
- Bài toán tìm cực trị đa thức với $\theta \in [0,1]$.

- Ta xét đa thức:

$$\mathrm{P}(\theta) = -\theta^6 + 5\theta^5 - 10\theta^4 + 10\theta^3 - 5\theta^2 + \theta$$

- $\mathrm{P}'(\theta) = -6\theta^5 + 25\theta^4 - 40\theta^3 + 30\theta^2 - 10\theta + 1.$
- Cho $\mathrm{P}'(\theta) = 0$. Tìm θ để $\mathrm{P}(\theta)$ đạt cực đại.

- Ta tìm được:

$$\theta = \frac{1}{6}$$

- Lưu ý: $\theta = \frac{1}{6} = \frac{1+1-1}{3+1+4-2} = \frac{1+\alpha-1}{3+\alpha+\beta-2}$
- Ta cũng có $\mathrm{P}(\theta|\mathrm{D}) \sim \mathrm{Beta}(2, 6) = \mathrm{Beta}\,(\alpha + 1, \beta + 2)$

- Như vậy về mặt tổng quát, trong trường hợp ta có phân phối tiên nghiệm là $\theta \sim \beta(\mathrm{m},\mathrm{n});\ \mathrm{D}|\theta \sim \mathrm{B}(\mathrm{a},\mathrm{b})$. Ước lượng bằng phương pháp Bayes với hàm Loss 0-1, ta được:

$$\hat{\theta} = \frac{\mathrm{m} - 1 + \mathrm{a}}{m + \mathrm{n} + a + \mathrm{b} - 2}$$

2. Sửa một vài bài tập:

- Bài tập: Chứng minh định lý ước lượng hàm loss bình phương trong trường hợp θ là một biến số thực.
- Định lý: Cho mẫu ngẫu nhiên $\mathrm{X}_1, \mathrm{X}_2, \dots, \mathrm{X}_\mathrm{n}$ với hàm mật độ xác suất $\mathrm{f}(\mathrm{X}|\theta)$ trong đó θ là tham số cần ước lượng. Nếu ta sử dụng hàm loss là hàm loss bình phương, thì ước lượng Bayes $\hat{\theta} = \mathrm{E}(\theta|\mathrm{x}_1, \mathrm{x}_2, \dots, \mathrm{x}_\mathrm{n})$.

Chứng minh:

- Xét hàm loss bình phương:

$$\mathrm{L}\left(\hat{\theta},\theta\right) = \left(\hat{\theta} - \theta\right)^2$$

- Khi đó ta có hàm loss kì vọng:

$$\mathrm{E}\left[\mathrm{L}\left(\hat{\theta},\theta\right)\right] = \int \left(\hat{\theta} - \theta\right)^2 \mathrm{p}(\theta|\mathrm{D})\mathrm{d}\theta)$$

- Theo như chứng minh định lý của bài trước, ta có:

$$\mathrm{E}\left[\mathrm{L}\left(\hat{\theta},\theta\right)\right] = \int \left(\hat{\theta}-\theta\right)^{2}\mathrm{p}(\theta|\mathrm{D})\mathrm{d}\theta$$

$$= \int \left(\left\{\hat{\theta}-\mathrm{E}(\theta|\mathrm{D})\right\}-\left\{\theta-\mathrm{E}(\theta|\mathrm{D})\right\}\right)^{2}\mathrm{p}(\theta|\mathrm{D})\mathrm{d}\theta \quad (1)$$

- (1) trở thành:

$$\int [\theta - E(\theta|\mathrm{D})]^2 \mathrm{p}(\theta|\mathrm{D})\mathrm{d}\theta + 2\Big(\hat{\theta}$$

$$- \mathrm{E}(\theta|\mathrm{D})\Big) \int \big(\theta - \mathrm{E}(\theta|\mathrm{D})\big)\mathrm{p}(\theta|\mathrm{D})\mathrm{d}\theta + \big(\mathrm{E}(\theta|\mathrm{D}) - \hat{\theta}\big)^2$$

- Ta đã biết:

$$2\left(\hat{\theta} - \mathrm{E}(\theta|\mathrm{D})\right)\int\left(\theta - \mathrm{E}(\theta|\mathrm{D})\right)\mathrm{p}(\theta|\mathrm{D})\mathrm{d}\theta$$

$$= 2\left(\hat{\theta} - \mathrm{E}(\theta|\mathrm{D})\right)\mathrm{E}[\theta - \mathrm{E}(\theta|\mathrm{D})] = 0$$

Với $\mathrm{E}[\theta - \mathrm{E}(\theta|\mathrm{D})] = 0, E[\theta - E(\theta)] = 0$.

- Và ta được:

$$\mathrm{E}\left[\mathrm{L}\left(\hat{\theta},\theta\right)\right]$$

$$= \int [\theta - E(\theta|\mathrm{D})]^2 \mathrm{p}(\theta|\mathrm{D})\mathrm{d}\theta + \left(\mathrm{E}(\theta|\mathrm{D}) - \hat{\theta}\right)^2$$

- Vì $\int[\theta - E(\theta|\mathrm{D})]^2 \mathrm{p}(\theta|\mathrm{D})\mathrm{d}\theta$ không phụ thuộc vào giá trị của tham số cần ước lượng $\hat{\theta}$, nên $\mathrm{E}\left[\mathrm{L}\left(\hat{\theta}, \theta\right)\right]$ đạt cực trị khi:

$$\left(\mathrm{E}(\theta|\mathrm{D}) - \hat{\theta}\right)^2 = 0$$

Hay $\mathrm{E}(\theta|\mathrm{D}) = \hat{\theta}$.

Bài tập (tiếp):

- Bài tập 2: Ước lượng tham số $\hat{\theta}$ trong trường hợp $\mathrm{D}|\theta$ có phân phối Poisson và hàm loss là một hàm bình phương tổn thất.

Lời giải bài 2:

- Giả sử ta có bộ dữ liệu $\mathrm{D} = (\mathrm{X}_1, \dots, \mathrm{X}_\mathrm{n})$ với:

$$\mathrm{p}(\mathrm{x_i}|\theta) = \frac{\theta^{\mathrm{x_i}} . \mathrm{e}^{-\theta}}{\mathrm{x_i}!}$$

- Ta xét hàm likelihood:

$$\mathrm{p}(\mathrm{D}|\theta) = \prod_{\mathrm{i}} \mathrm{p}(\mathrm{x_i}|\theta) = \frac{\theta^{\sum_\mathrm{i} \mathrm{x_i}} . \mathrm{e}^{-\mathrm{n}\theta}}{\prod_\mathrm{i} \mathrm{x_i}!} = \frac{\theta^{n\bar{x}} . \mathrm{e}^{-\mathrm{n}\theta}}{\prod_\mathrm{i} \mathrm{x_i}!}$$

- Ta cho phân phối tiên nghiệm là $\theta \sim \mathrm{Gamma}(\alpha, \beta)$ và:

$$\mathrm{p}(\theta) = \frac{\theta^{\alpha-1}\beta^{\alpha}.\,\mathrm{e}^{-\beta\theta}}{\Gamma(\alpha)}$$

- Từ đây, ta có xác suất hợp:

$$\mathrm{P}(\mathrm{X}_1, \dots, \mathrm{X}_\mathrm{n}; \theta) = \mathrm{P}(\mathrm{D}|\theta).\,\mathrm{P}(\theta) = \frac{\beta^{\alpha}\theta^{\mathrm{n}\bar{\mathrm{x}}+\alpha-1}.\,\mathrm{e}^{-(\mathrm{n}+\beta)\theta}}{\prod_\mathrm{i} \mathrm{x}_\mathrm{i}!\,.\,\Gamma(\alpha)}$$

- Sử dụng xác suất biên, ta có:

$$\mathrm{P(D)} = \int_0^{+\infty} \mathrm{P(D}, \theta)\, d\theta$$

$$= \int_0^{+\infty} \frac{\beta^{\alpha} \theta^{\mathrm{n}\overline{\mathrm{x}}+\alpha-1} .\, \mathrm{e}^{-(\mathrm{n}+\beta)\theta}}{\prod_{\mathrm{i}} \mathrm{x_i}!\, .\, \Gamma(\alpha)} d\theta$$

- Tính toán với tích phân, ta có:

$$\mathrm{P(D)} = \frac{\beta^{\alpha}.\ \Gamma(\mathrm{n}\overline{\mathrm{x}} + \alpha)}{\prod_{\mathrm{i}} \mathrm{x_i}!.\ (\mathrm{n} + \beta)^{\mathrm{n}\overline{\mathrm{x}} + \alpha}}$$

- Xét công thức:

$$\mathrm{P}(\theta|\mathrm{D}) = \frac{\mathrm{P}(\mathrm{D}, \theta)}{\mathrm{P(D)}}$$

- Ta được:

$$P(\theta|D) = \frac{(n+\beta)^{n\overline{x}+\alpha}}{\Gamma(n\overline{x}+\alpha)} \theta^{n\overline{x}+\alpha-1}.\ e^{-\theta(n+\beta)}$$

- Vậy $\theta|D \sim \mathrm{Gamma}(n\overline{x}+\alpha, n+\beta)$

- Như vậy, áp dụng định lý ước lượng Bayes dành cho hàm loss bình phương, ta được tham số $\hat{\theta}$:

$$\hat{\theta} = \mathrm{E}(\theta|\mathrm{D}) = \frac{\alpha'}{\beta'} = \frac{\mathrm{n}\bar{\mathrm{x}} + \alpha}{\mathrm{n} + \beta}$$

Bài tập: cho $D = (X_1 = 3, X_2 = 5, X_3 = 1)$ và $D|\theta$ có phân phối Poisson. Biết $\theta \sim Gamma(3,2)$. Tính $\hat{\theta}$.

Bài tập (tiếp)

- Bài tập 3: Ước lượng tham số $\hat{\theta}$ trong trường hợp $\mathrm{D}|\theta$ có phân phối chuẩn và hàm loss là một hàm bình phương tổn thất.

Lời giải:

- Giả sử ta có bộ dữ liệu $\mathrm{D} = (\mathrm{X}_1, \dots, \mathrm{X}_\mathrm{n})$ với $\mathrm{X}_\mathrm{i} \sim \mathrm{N}(\theta, 1)$:

$$\mathrm{p}(\mathrm{x}_\mathrm{i}|\theta) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}(\mathrm{x}-\mu)^2}$$

- Khi đó, hàm likelihood sẽ là:

$$\mathrm{p}(\mathrm{D}|\theta) = \frac{1}{(2\pi)^{\frac{\mathrm{n}}{2}}} \mathrm{e}^{\sum_{\mathrm{i}=1}^{\mathrm{n}}\left(-\frac{1}{2}(\mathrm{x}_\mathrm{i}-\mu)^2\right)}$$

- Ta xét xác suất tiên nghiệm $\theta \sim \mathrm{N}(0,1)$:

$$\mathrm{P}(\theta) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\theta^2}$$

- Từ đó ta được phân phối hợp:

$$\mathrm{P}(\theta, \mathrm{D}) = \frac{1}{(2\pi)^{\frac{n+1}{2}}} \mathrm{e}^{-\frac{1}{2}\left(\sum_{i=1}^{n} x_i^2 - 2\theta n\bar{x} + (n+1)\theta^2\right)}$$

- Xác suất biên của D sẽ là:

$$P(D) = \int_{-\infty}^{+\infty} P(D, \theta) d\theta$$

$$= \frac{1}{\sqrt{n+1}(2\pi)^{\frac{n+1}{2}}} e^{-\frac{1}{2}\left(\sum_{i=1}^{n} x_i^2 - \frac{n^2(\overline{x})^2}{n+1}\right)}$$

- Như vậy ta được phân phối hậu nghiệm:

$$\mathrm{P}(\theta|\mathrm{D}) = \frac{\mathrm{P}(\theta,\mathrm{D})}{\mathrm{P}(\mathrm{D})}$$

$$= \frac{\sqrt{\mathrm{n}+1}}{\sqrt{2\pi}}\,\mathrm{e}^{\frac{-(n+1)}{2}\left(\theta-\frac{\mathrm{n}\bar{\mathrm{x}}}{\mathrm{n}+1}\right)^2}$$

- Ta sẽ áp dụng định lý ước lượng tham số cho hàm loss bình phương để tính $\hat{\theta}$.
- Ta có:

$$\hat{\theta} = \mathrm{E}(\theta|\mathrm{D}) = \int_{-\infty}^{+\infty} \theta \mathrm{P}(\theta|\mathrm{D}) \mathrm{d}\theta = \frac{\mathrm{n}\overline{\mathrm{x}}}{\mathrm{n}+1}$$

- Lưu ý: Ta có thể kết luận thẳng là $\hat{\theta} = \frac{\mathrm{n}\bar{\mathrm{x}}}{\mathrm{n}+1}$, vì nó có phân phối xác suất:

$$\theta|\mathrm{D} \sim \mathrm{N}\left(\frac{\mathrm{n}\bar{\mathrm{x}}}{\mathrm{n}+1}, \sqrt{\frac{1}{\mathrm{n}+1}}\right)$$

Khi đó $\hat{\theta} = \mathrm{E} = \frac{\mathrm{n}\bar{\mathrm{x}}}{\mathrm{n}+1}$