# Bài 1:

Chú ý: phân phối Gamma có hai dạng là Shape - rate và shape - scale.

Shape - rate: Gamma$(\alpha, \beta)$. Hàm mật độ là $p(x;\alpha,\beta) = \dfrac{x^{\alpha-1} e^{-\beta x} \beta^{\alpha}}{\Gamma(\alpha)}$

Shape - scale: Gamma$(k, \theta)$. Hàm mật độ là $p(x; k, \theta) = \dfrac{x^{k-1} e^{-\frac{x}{\theta}} . (\theta^{-1})^{k}}{\Gamma(k)}$

So sánh: $\alpha = k$, $\beta = \dfrac{1}{\theta}$. Như vậy ta hoàn toàn có thể dùng cả hai dạng. Tuy nhiên cần để ý đến trung bình và phương sai theo từng dạng.

Shape-scale: $E = k\theta$, $Var = k\theta^2$

Shape-rate: $E = \dfrac{\alpha}{\beta}$, $Var = \dfrac{\alpha}{\beta^2}$

Để thuận tiện cho về sau, chúng ta sẽ coi p.p.s.x Gamma $(\alpha, \beta)$ hoặc Gamma $(n, \theta)$ có dạng Shape - rate.

a) Ước lượng theo maximum likelihood.

$$P(X_1, X_2, \ldots, X_n) = \prod_{i=1}^{n} P(X_i|\theta) = \prod_{i=1}^{n} \frac{e^{-\theta}\theta^{x_i}}{x_i!} = \frac{e^{-n\theta}\theta^{\sum_{i=1}^{n} x_i}}{\prod x_i!}$$

$$\underset{\theta}{\text{Arg max}}\ P(X_1, \ldots, X_n) = \text{Arg max}\ e^{-n\theta}\theta^{\sum_{i=1}^{n} x_i}$$

Tìm cực trị của $e^{-n\theta}\theta^{\sum_{i=1}^{n} x_i}$:

$$\underset{\theta}{\arg\max}\ e^{-n\theta}\theta^{\sum_{i=1}^{n} x_i} = \underset{\theta}{\arg\max}\ \ln(e^{-n\theta}\theta^{\sum_{i=1}^{n} x_i}) = \underset{\theta}{\arg\max} \underbrace{-n\theta + \ln \theta^{\sum_{i=1}^{n} x_i}}_{f(\theta)}$$

$$f'(\theta) = -n + \sum_{i=1}^{n} x_i \cdot \frac{1}{\theta}$$

$$f'(\theta) = 0 \Leftrightarrow -n + \frac{1}{\theta}\sum x_i = 0 \Leftrightarrow \theta = \frac{\sum x_i}{n}$$

$$\Rightarrow \hat{\theta} = \frac{\sum x_i}{n}$$

b) Tìm P.P.S.X ~~tiên~~ hậu nghiệm:

Ta sẽ tìm $\theta | X_1, \dots, X_n$ biết $\theta \sim Gamma(\alpha, \beta)$

Ta có: $P(\theta | X_1, \dots, X_n) = \dfrac{P(X_1, \dots, X_n | \theta) . P(\theta)}{P(X_1, \dots, X_n)}$

$$\propto P(X_1, \dots, X_n | \theta) . P(\theta)$$

$$\propto e^{-n\theta} \theta^{\sum_{i=1}^{n} x_i} . \theta^{\alpha - 1} e^{-\beta . \theta}$$

$$\propto \theta^{\sum_{i=1}^{n} x_i + \alpha - 1} . e^{-(n+\beta)\theta}$$

Như vậy ta được $\theta | X_1, \dots, X_n \sim Gamma\left(\sum_{i=1}^{n} x_i + \alpha, n + \beta\right)$ (theo dạng shape-rate)

$E(\theta | X_1, \dots, X_n) = \dfrac{\sum_{i=1}^{n} x_i + \alpha}{n + \beta}$, $Var(\theta | X_1, \dots, X_n) = \dfrac{\sum_{i=1}^{n} x_i + \alpha}{(n+\beta)^2}$

2

a) Theo câu 1a, ta có:

$$\hat{\theta} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{\sum_{i=1}^{100} i}{100} = 50,5.$$

P.p.h.n có dạng:

$$\theta | X_1, \dots, X_n \sim Gamma\left(\sum_{i=1}^{n} x_i + \alpha, n + \beta\right) \sim Gamma(5050 + 50, 100 + 100)$$

$$\sim Gamma(5100, 200)$$

b) Cập nhật thêm 5 mẫu vào data ta cũ đồng nghĩa với việc tăng n từ 100 lên 105 với $x_i = i$, $i = \overline{101, 105}$. Như vậy cũng tương tự câu a), ta có:

$$\hat{\theta} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{\sum_{i=1}^{105} x_i}{105} = 53$$

Nx: Sau khi thêm mẫu vào thì giá trị $\hat{\theta}$ tăng lên. ~~Theo định luật số lớn~~

3.

Tính p.p.s.x hậu nghiệm:

$$p(X|\theta) = C_5^x \theta^x (1-\theta)^{5-x} \quad (\text{likelihood})$$

$$p(\theta) = \frac{\Gamma(1+9)}{\Gamma(1)\Gamma(9)} \theta^{1-1}(1-\theta)^{9-1} = 9(1-\theta)^8$$

$$p(\theta|x) \propto p(x|\theta).p(\theta) \propto \theta^x (1-\theta)^{13-x}.$$

$$\Rightarrow \theta|x \sim Beta(x+1, 14-x).$$

a) Ước lượng Bayes theo hàm loss $L = (\theta - a)^2$ là:

$\hat{\theta} = \arg\max E[L(\hat{\theta},\theta)] = E(\theta|x)$ (định lý trang 37 - Slide bài 4).

Mà ta có. $E(\theta|x) = \frac{\alpha}{\alpha+B}$ (trung bình của phân phối Beta)

$$= \frac{x+1}{x+1+14-x} = \frac{x+1}{15}$$

b) Ước lượng Bayes theo hàm loss $L = |\theta - a|$ là:

$\hat{\theta} = median(\theta|X)$ (định lý trang 12 - bài 4).

$median(\theta|x)$ là giá trị m sao cho $F(\theta|x) = \frac{1}{2}$

Tức là:

$$\int_0^m p(\theta|x)\, d\theta = \frac{1}{2}$$

Vì x là một giá trị cố định, ta có thể cho $x=0$, khi đó bài toán trở thành tìm m sao cho

$$\int_0^m p(\theta|x=0)\, d\theta = \frac{1}{2}$$

$$\Leftrightarrow \int_0^m \frac{\Gamma(15)}{\Gamma(1)\Gamma(14)} \theta^{1-1}(1-\theta)^{14-1} d\theta = \frac{1}{2}$$

$$\Leftrightarrow \int_0^m 14(1-\theta)^{13} d\theta = \frac{1}{2}$$

$$\Leftrightarrow 14.\left(\frac{-1}{14}\right)(1-\theta)^{14}\Big|_0^m = \frac{1}{2}$$

$$\Leftrightarrow 1-(1-m)^{14} = \frac{1}{2} \Leftrightarrow m = 1-\left(\frac{1}{2}\right)^{\frac{1}{14}} = 1-2^{-\frac{1}{14}}$$

4.

Tính hàm hợp xác suất:

$$P(x_1, \theta) = P(x|\theta).P(\theta)$$

$$= \begin{cases} 0{,}2 \cdot 0{,}6 \cdot e^{-\frac{\theta}{1000}} = 0{,}12 \;;\; \theta = 0 \\ 0{,}6 \cdot e^{-\frac{\theta}{1000}} \cdot \frac{0{,}1}{3000} \cdot e^{-\frac{\theta}{3000}} = \frac{0{,}06}{3000} \cdot e^{-\frac{\theta}{750}} \end{cases}$$

$$m(x) = \int_{-\infty}^{\infty} P(x,\theta)\,d\theta = 0{,}12 + \int_0^{\infty} \frac{0{,}06}{3000} \cdot e^{-\frac{\theta}{750}}\,d\theta$$

$$= 0{,}12 + \frac{0{,}06}{3000} * \int_0^{\infty} e^{-\frac{\theta}{750}}\,d\theta$$

Tính $\int_0^{\infty} e^{-\frac{\theta}{750}}\,d\theta$

$\int e^{-\frac{\theta}{750}}\,d\theta = -750\,e^{-\frac{\theta}{750}}$, vậy nên

$$\int_0^{\infty} e^{-\frac{\theta}{750}}\,d\theta = \lim_{t\to\infty} \int_0^t e^{-\frac{\theta}{750}}\,d\theta = \lim_{t\to\infty} -750\,e^{-\frac{\theta}{750}}\Big|_0^t$$

$$= \lim_{t\to\infty} -750\,e^{-\frac{t}{750}} + 750$$

$$= 750$$

$$\Rightarrow m(x) = 0{,}12 + \frac{0{,}06}{3000} \cdot 750 = 0{,}135$$

Ta có p.p.x.s hậu nghiệm:

$$P(\theta|x) = \frac{P(x,\theta)}{P(x)} = \begin{cases} \frac{0{,}12}{0{,}135} \;;\; \theta = 0 \\ \dfrac{\frac{0{,}06}{3000} \cdot e^{-\frac{\theta}{750}}}{0{,}135} \end{cases}$$

$$= \begin{cases} \frac{8}{9} \;;\; \theta = 0 \\ \frac{4}{9} \cdot \frac{e^{-\frac{\theta}{750}}}{3000} \;,\; \theta > 0. \end{cases}$$

Bài 4 (tiếp):

Tính Bayesian loss cho từng quyết định:

$$E^{p(\theta|x)}[L(\theta,a_1)] = \frac{8}{9}\cdot 100000 + \int_0^{\infty}(10000-2\theta)\,\frac{4}{9}\cdot\frac{e^{-\frac{\theta}{750}}}{30000}\,d\theta$$

$$= \frac{800000}{9} + \frac{98500}{9} = 99833{,}3.$$

chú ý: $E^{p(\theta|x)}[L(\theta,a_1)] = \int_{-\infty}^{\infty} L(\theta,a_1).\,p(\theta|x)\,d\theta$.

Tương tự, ta có: $E^{p(\theta|x)}[L(\theta,a_2)] = 46575$, $E^{p(\theta|x)}[L(\theta,a_3)] = 6641{,}6$

So sánh ta có: $E^{p(\theta|x)}[L(\theta,a_1)] > E^{p(\theta|x)}[L(\theta,a_2)] > E^{p(\theta|x)}[L(\theta,a_3)]$

Chọn $a_3$.

Bài 5:

Cách 1: Theo tính chất của phân phối ~~phổ~~ gamma và liên hệ với phân phối chi bình phương, ta có:

$X \sim Gamma\left(\frac{v}{2}, \frac{1}{2}\right)$ (theo ~~sph~~ shape-rate) nếu $X \sim \chi^2(v)$

Trong TH này ta có $v = 24$ nên $X \sim Gamma\left(12, \frac{1}{2}\right)$.

Theo câu 1, ta có:

$$p(\theta|x_1,\dots,x_{12}) \propto p(x_1,\dots,x_{12}|\theta).\,p(\theta)$$

$$\propto e^{-12\theta}.\,\theta^{\sum x_i}.\,\theta^{12-1}.\,e^{-\frac{1}{2}\theta}$$

$$= e^{-(12+\frac{1}{2})\theta}.\,\theta^{\sum x_i + 12 - 1}$$

Với $\sum x_i = 11+21+35+40+60+68+61+13+23+23+21+41 = 447$

$\Rightarrow \theta|x_1,\dots,x_{12} \sim Gamma(447+12,\ 12+\frac{1}{2}) = Gamma(459,\ 12{,}5)$.

Cách 2:

Theo slide bài 3.2, trang 63, ta có:

$$\theta|x_1,\dots,x_{12} \sim (2n)^{-1}\chi^2_v$$

Ta có $v=24$ và $n=12$, vậy nên $\theta|x_1,\dots,x_{12} \sim (24)^{-1}.\,\chi^2_{24}$

Bài 6:

a) Likelihood: $x_i | \theta \sim N(\theta, \phi)$ $\phi = 2,5$

$\bar{x} | \theta \sim N(\theta, \frac{\phi}{n}) = N(\theta, 0,5)$ và $\bar{x} = \frac{\sum_{i=1}^{5} x_i}{n} = 17,2$

Để kiểm tra ~~1 trang 2~~ các giả thuyết:

$H_0: \theta = 19$ ($\theta$ được gọi là $\theta_0$ nếu $\theta = 19$)

$H_1: \theta \neq 19$

Ta sẽ tính $B = \frac{\frac{P_0}{P_1}}{\frac{\pi_0}{\pi_1}} = \frac{p(x|\theta_0)}{p(x)} = \frac{p(x_1, \dots, x_5 | \theta = 19)}{p(x)}$

Xác suất ~~hậu~~ tiên nghiệm: $\theta \sim N(\mu = 19, \psi = 3)$

Theo slide 42 ~~trang~~ bài 5, ta có:

$$B = \frac{\left\{\frac{2\pi\phi}{n}\right\}^{-\frac{1}{2}} e^{-\frac{\frac{1}{2}(\bar{x}-\theta_0)^2}{\frac{\phi}{n}}}}{\left\{2\pi(\psi + \frac{\phi}{n})\right\}^{-\frac{1}{2}} e^{-\frac{\frac{1}{2}(\bar{x}-\theta_0)^2}{\psi + \frac{\phi}{n}}}}$$

Ta có: $\bar{x} = 17,2$, $\frac{\phi}{n} = 0,5$, $\theta_0 = 19$, $\psi = 3$, ta được:

$$B = \frac{\sqrt{2\pi . 0,5} . e^{-1,8^2}}{\sqrt{2\pi . 3,5} . e^{-\frac{\frac{1}{2} . 1,8^2}{3,5}}} \simeq 0,163 < 1$$

Vậy ta bác bỏ $H_0$

b) Ta có, tìm phân phối hậu nghiệm:

$$\theta \mid x_1, \dots, x_n \sim N(\theta_1, \mu_1) \quad \text{(theo slide 16, bài 3.2)}$$

$$\cancel{\mu_1 = \frac{1}{\phi^{-1} + n\psi^{-1}} = \frac{1}{0,4 + \frac{5}{3}} = 0,48}$$

$$\cancel{\theta_1 = \mu_1 \cdot \left(\frac{19}{3} + \frac{\sum_{i=1}^{5} x_i}{2,5}\right) = 0,48 \cdot \left(\frac{19,2}{3} + 34,4\right) \simeq 9,}$$

$$\mu_1 = \frac{1}{\phi^{-1} + n\psi^{-1}} = \frac{1}{0,4 + 5 \cdot \frac{1}{3}} \simeq 0,48$$

$$\theta_1 = \mu_1 \left(\frac{\mu}{\psi} + \frac{\bar{x}}{\frac{\phi}{n}}\right) = 0,48 \left(\frac{19}{3} + \frac{17,2}{0,5}\right) \simeq 19,55$$

Như vậy: $\theta \mid x_1, \dots, x_n \sim N(19,55;\ 0,48)$

Để kiểm $H_0 : \theta \leq 19$ và $H_1 : \theta > 19$, ta có:

$$\alpha_0 = P(\theta \leq 19 \mid X_1, \dots, X_n) = P\left(\frac{\theta - 19,55}{\sqrt{0,48}} \leq \frac{19 - 19,55}{\sqrt{0,48}} \mid x_1, \dots, x_n\right)$$

$$\simeq P(Z \leq -0,797)$$

$$\simeq 0,21$$

~~Suy ra ta bác bỏ $\alpha_0$.~~

$$\alpha_1 = P(\theta > 19 \mid X_1, \dots, X_n) = 1 - \alpha_0 \approx 0,79$$

$$\Rightarrow \frac{\alpha_0}{\alpha_1} < 1 \Rightarrow \text{bác bỏ } H_0.$$