AST_EX_4

roi hezkiyahu

28 12 2022

Q1

Question 1.

Assume that a distribution of Y belongs to a one-parametric parametric family of distributions f(y|θ), where a priori θ ~ π(θ). After getting the first observation y_1 we update the distribution of θ by calculating its posterior distribution $p(\theta|y_1)$ and use it as a *prior* distribution for θ before we observe y_2. Having y_2, we update the distribution of θ again by deriving its posterior distribution. Show that one will get the same posterior distribution of θ if instead of caclulating it sequentially, he/she would use the whole data (i.e. y_1 and y_2) and the original prior π(θ).

$$\text{let } y_1 \text{ be a sample from } Y \sim f(y|\theta); \theta \sim \pi(\theta)$$
$$p(\theta|y_1) = \frac{f(y_1|\theta)\cdot\pi(\theta)}{f(y_1)} := \pi_1(\theta)$$
$$\text{assume } \pi_1(\theta) \text{ is the prior for sampling } y_2 \text{ we get:}$$
$$p(\theta|y_2) = \frac{f(y_2|\theta)\cdot\pi_1(\theta)}{f(y_2)} = \frac{f(y_2|\theta)}{f(y_2)}\frac{f(y_1|\theta)\cdot\pi(\theta)}{f(y_1)} := \hat{\pi}(\theta)$$
$$p(\theta|\underline{y}) = \frac{f(\underline{y}|\theta)\cdot\pi(\theta)}{f(\underline{y})} = \frac{f(y_2|\theta)}{f(y_2)}\frac{f(y_1|\theta)\cdot\pi(\theta)}{f(y_1)}$$

Q2

Question 2.

Let $Y_1,\dots,Y_n \sim f_\theta(y)$, where f_θ belongs to the exponential family, i.e. $supp f_\theta$ does not depend on θ and $f_\theta(y) = \exp\{c(\theta)T(y) + d(\theta) + S(y)\}$.

1. Show that the conjugate prior family for θ is $\pi_{a,b}(\theta) \propto \exp\{ac(\theta) + bd(\theta)\}$.
2. Find a conjugate prior for the parameter λ of the Poisson distribution Pois(λ).

Q2.1

$$\pi(\theta|y) = \frac{f(\theta|y)\pi(\theta)}{f(y)} \propto L(\theta;y)\pi(\theta)$$
$$L(\theta;y) = \Pi_i exp\{c(\theta)T(y_i) + d(\theta) + S(y_i)\} = exp\{c(\theta)\sum_i T(y_i) + bd(\theta)\}\cdot exp(\sum_i S(y_i)) \propto exp\{T'(y)c(\theta) + nd(\theta)\}$$
$$\text{plug it back in and we get:}$$
$$\pi(\theta|y) \propto L(\theta;y)\pi(\theta) \propto exp\{T'(y)c(\theta) + nd(\theta)\}\cdot exp\{ac(\theta) + bd(\theta)\} = exp\{(T'(y) + a)c(\theta) + (n+b)d(\theta)\}$$
$$\text{thus we get that the conjecture prior family is the exponenital family}$$

Q2.2

$$\text{for the possion family the:}$$
$$T(y) = \sum_i y_i; \quad c(\lambda) = ln(\lambda); \quad d(\lambda) = n\lambda$$
$$\lambda \sim exp(ln(\theta) + n\theta) = gamma(1, ln(\theta) + n\theta)$$
$$\text{lets show it's conjecture to the poison:}$$
$$\pi(\lambda|y) \propto L(\lambda;y)\pi(\lambda) = e^{-n\lambda}\cdot\frac{\lambda^{T(y)}}{\Pi_i y_i}(ln(\theta) + n\theta)e^{-(ln(\theta)+n\theta)\lambda} \propto$$
$$\propto \lambda^{T(y)}e^{-(ln(\theta)+n\theta+n)\lambda} \Rightarrow \lambda|y \sim Gamma(T(y)+1, (ln(\theta) + n\theta + n))$$

Q3

Question 3.

Let $Y_1, \ldots, Y_n \sim Pois(\lambda)$.

1. Assume *a priori* that $\lambda \sim \exp(\theta)$. Is it a conjugate prior for Poisson data?
2. Estimate λ and $p = P(Y = 0)$ w.r.t. the quadratic loss.
3. Assume now that θ is also unknown and find its empirical Bayes estimator. What are the resulting estimators for λ and p?

Q3.1

the answer is yes, this is straightforward from the last question
poisson, gamma, and the exponential distribution all belong to he exponential family

Q3.2

$$\lambda|y \sim Gamma(T(y)+1, \theta+n)$$
$$\hat{\lambda} = E_{\lambda|y}[\lambda|y] = \frac{T(y)+1}{\theta+n}$$
$$\hat{p} = \hat{P}(Y=0) = exp\{-\frac{T(y)+1}{\theta+n}\}$$

Q3.3

$$L(\lambda; y) = \Pi_i f(y_i|\lambda) = e^{-n\lambda} \cdot \frac{\lambda^{T(y)}}{\Pi_i y_i}$$
$$\hat{\lambda}_{MLE} = \bar{y}$$
$$\bar{y} = \frac{T(y)+1}{\theta+n} \Rightarrow \hat{\theta}_{MLE} = \frac{T(y)+1}{\bar{y}} - n = \frac{1}{\bar{y}}$$

Q4

Question 4.

Suppose that $Y_1, \ldots, Y_n \sim \exp(\theta)$, where $EY = 1/\theta$.

1. Find a noninformative prior for θ according to the Jefrreys' rule and the corresponding posterior distribution.
2. Estimate θ w.r.t. the quadratic error and compare the resulting Bayesian estimator with the MLE.
3. Repeat the previous paragraph for estimating $\mu = EY$ and $p = P(Y \geq a)$.
4. In addition to the sample of Y's, we have another independent sample $X_1, \ldots, X_n \sim \exp(\phi)$, $EX = 1/\phi$ and again, we use the noninformative prior for ϕ. We are interested in the ratio θ/ϕ. Find its posterior distribution, the posterior mean and compare it with the MLE for this ratio.

(hint: recall some of basic distributions you know like χ^2, F, etc.)

Q4.1

$$\theta \sim \pi(\theta)$$
from jefrrey's rule:
$$\pi(\theta) \propto \sqrt{I^*(\theta)} = \sqrt{E(-l(\theta; y)'')} = \sqrt{-E(-\frac{1}{\theta^2})} = \frac{1}{\theta} = \theta^{-1}$$
we derived the noninformative prior now lets calculate the posterior:
$$\pi(\theta|y) \propto L(\theta; y)\pi(\theta) = \theta^n e^{-n\bar{y}\theta}\theta^{-1} = \theta^{n-1}e^{-n\bar{y}\theta} \Rightarrow \theta|y \sim Gamma(n, n\bar{y})$$

Q4.2

$$\theta|y \sim Gamma(n, n\bar{y})$$
$$\hat{\theta} = E_{\theta|y}[\theta|y] = \frac{n}{n\bar{y}} = \frac{1}{\bar{y}}$$
$$\hat{\theta}_{MLE} = \frac{1}{\bar{y}} = \hat{\theta}$$

Q4.3

$$\hat{\mu} = E(\frac{1}{\theta}|y) = \int_0^\infty \frac{1}{\theta}\theta^{n-1}e^{n\bar{y}\theta}\frac{(n\bar{y})^n}{\Gamma(n)}d\theta = \int_0^\infty \theta^{n-1-1}e^{n\bar{y}\theta}\frac{(n\bar{y})^n}{\Gamma(n)}\frac{\Gamma(n-1)}{\Gamma(n-1)} =$$
$$\frac{\Gamma(n-1)}{\Gamma(n)}n\bar{y}\int_0^\infty \theta^{n-1-1}e^{n\bar{y}\theta}\frac{(n\bar{y})^{n-1}}{\Gamma(n-1)} = \frac{n\bar{y}}{n-1}$$
$$p = P(Y \geq a) = e^{-\theta a}$$
$$\hat{p} = E(e^{-\theta a}|y) = \int_0^\infty e^{-\theta a}\theta^{n-1}e^{n\bar{y}\theta}\frac{(n\bar{y})^n}{\Gamma(n)}d\theta = \int_0^\infty \theta^{n-1}e^{(n\bar{y}-a)\theta}\frac{(n\bar{y})^n}{\Gamma(n)}d\theta = \int_0^\infty \theta^{n-1}e^{(n\bar{y}-a)\theta}\frac{(n\bar{y})^n}{\Gamma(n)}\frac{(n\bar{y}-a)^n}{(n\bar{y}-a)^n}d\theta =$$
$$= \frac{(n\bar{y})^n}{(n\bar{y}-a)^n}\int_0^\infty \theta^{n-1}e^{(n\bar{y}-a)\theta}\frac{(n\bar{y}-a)^n}{\Gamma(n)}d\theta = \frac{(n\bar{y})^n}{(n\bar{y}-a)^n}$$
$$\hat{\mu}_{MLE} = \frac{1}{\hat{\theta}_{MLE}} = \bar{y} \neq \hat{\mu}$$
$$\hat{p}_{MLE} = e^{-\hat{\theta}_{MLE}a} = e^{-\bar{y}a} \neq \hat{p}$$

Q4.4

$$\text{if } X_1, X_2 \sim Gamma(\alpha_i, \beta_i) \text{ then } \frac{\beta_2\alpha_2X_1}{\beta_1\alpha_1X_2} \sim F(2\alpha_1, 2\alpha_2)$$

using the idenity above and the fact that are paramaters are gamma distributed we can derive:

$$\frac{nn\bar{Y}\theta}{nn\bar{X}\phi}|(X,Y) \sim F(2n, 2n) \Rightarrow \frac{\theta}{\phi}|(X,Y) \sim \frac{\bar{X}}{\bar{Y}}F(2n, 2n)$$
$$\text{the posterior mean is simply : } \frac{\bar{X}}{\bar{Y}}\frac{2n}{2n-2} = \frac{\bar{X}}{\bar{Y}}\frac{n}{n-1}$$
$$\text{the MLE for this ratio is: } \frac{\bar{X}}{\bar{Y}}$$

the MLE the the posterior mean are different but as n goes to infinity they become closer

Q5

Question 5.

Three friends, Optimist, Realist and Pessimist, are going to participate in a certain kind of gambling game in casino when they do not know the probability of win *p*. Motivated by an exciting course in Bayesian statistics, they decided to apply Bayesian analysis. Optimist chose a prior on *p* from the conjugate family of distributions. In addition, he believes that the chances are 50%-50% (i.e. E(p)=1/2) with Var(p)=1/36. Realist chose a uniform prior U[0,1].

1. Show that both priors belong to the family of Beta distributions and find their parameters for the priors of Optimist and Realist. (hint: $E\left(Beta(\alpha,\beta)\right)=\frac{\alpha}{\alpha+\beta},\ Var\left(Beta(\alpha,\beta)\right)=\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$)
2. Pessimist does not have any prior beliefs and decides to choose the noninformative prior according to the Jeffreys' rule. What is his prior? Does it also belong to the Beta family?
3. Being poor students in Statistics, Optimist, Realist and Pessimist do not have enough money to gamble separately, so they decided to play together. They played the game 25 times and won 12 times. What is the posterior distribution of each one of them?
4. Find the corresponding posterior means and calculate the corresponding 95% Bayesian credible intervals for *p*. (hint: use the fact that if p~Beta(α,β) and ρ=p/(1-p) is the odds, then (β/α)ρ ~ $F_{2\alpha,2\beta}$ - those who do not believe it, can (should!) easily verify it!)
5. Three friends told their classmate Skeptic about the exciting Bayesian analysis each one of them has done. However, Skeptic is, naturally, very skeptical about Bayesian approach and does not belive in any priors - his credo is "*In G-d we trust... All others bring data*". So he decided to perform a "classical" (non-Bayesian) analysis of the same data. What is his estimate and the 95%-confidence interval for *p*? Compare the results and comment them briefly.

Q5.1

$$E(p_{opt})=1/2=\frac{\alpha}{\alpha+\beta}\Rightarrow\alpha=\beta$$
$$V(p_{opt})=1/36=\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}=\frac{\alpha^2}{(2\alpha)^2(2\alpha+1)}=\frac{1}{8\alpha+4}\Rightarrow\alpha=\beta=4$$
$$\pi(p_{opt})\sim Gamma(4,4)$$
$$U(0,1)=Beta(1,1)\text{ thus the prior for the Realist also belong to the beta distribution}$$

Q5.2

$$\pi(p_{pes})\propto\sqrt{I^*(p)}=\sqrt{E(-l(p;y)'')}=\sqrt{E(\frac{y_i}{p^2}+\frac{1-y_i}{(1-p)^2})}=\sqrt{(\frac{p}{p^2}+\frac{(1-p)}{(1-p)^2})}=\sqrt{(\frac{1}{p}+\frac{1}{1-p})}=$$
$$=\sqrt{\frac{1}{p(1-p)}}=p^{-0.5}(1-p)^{-0.5}\Rightarrow\pi(p_{pes})\sim Beta(0.5,0.5)$$

Q5.3

$$\pi(p_{opt}|y)\propto L(p;y)\pi(p_{opt})=p^{12}(1-p)^{13}p^3(1-p)^3\Rightarrow\pi(p_{opt}|y)\sim Beta(16,17)$$
$$\pi(p_{rel}|y)\propto L(p;y)\pi(p_{rel})=p^{12}(1-p)^{13}p^0(1-p)^0\Rightarrow\pi(p_{rel}|y)\sim Beta(13,14)$$
$$\pi(p_{pes}|y)\propto L(p;y)\pi(p_{pes})=p^{12}(1-p)^{13}p^{-0.5}(1-p)^{-0.5}\Rightarrow\pi(p_{pes}|y)\sim Beta(12.5,13.5)$$

Q5.4

$$\text{before we start lets derive a general form for the creddibal sets :}$$
$$\rho \in (L,U) \iff \frac{p}{1-p} \in (L,U) \iff \frac{1-p}{p} \in (1/U, 1/L) \iff \frac{1}{p} \in (\frac{U+1}{U}, \frac{L+1}{L}) \iff p \in (\frac{L}{L+1}, \frac{U}{U+1})$$
$$\rho \sim \frac{\alpha}{\beta} F_{2\alpha,2\beta} \Rightarrow (L,U) = (\frac{\alpha}{\beta} F_{2\alpha,2\beta,0.025}, \frac{\alpha}{\beta} F_{2\alpha,2\beta,0.975})$$
$$\text{optimist:}$$
$$\hat{p}_{opt} = \frac{16}{16+17} = 0.484$$
$$\alpha/\beta = 16/17$$
$$F_{32,34,0.025} = 0.497, \quad F_{32,34,0.975} = 1.996$$
$$P(p_{opt} \in (0.318, 0.651)) = 0.95$$
$$\text{realist:}$$
$$\hat{p}_{rel} = \frac{13}{13+14} = 0.481$$
$$\alpha/\beta = 13/14$$
$$F_{26,28,0.46} = 0.459, \quad F_{26,28,0.975} = 2.15$$
$$P(p_{rel} \in (0.298, 0.666)) = 0.95$$
$$\text{pessist:}$$
$$\hat{p}_{pes} = \frac{12.5}{12.5+13.5} = 0.4807$$
$$\alpha/beta = 12.5/13.5$$
$$F_{25,27,0.46} = 0.452, \quad F_{25,27,0.975} = 2.182$$
$$P(p_{pes} \in (0.295, 0.669)) = 0.95$$

Q5.5

$$\hat{p}_{MLE} = 12/25 = 0.48$$
$$\text{from CLT: } \hat{p}_{MLE} \dot{\sim} N(p, \frac{p(1-p)}{n})$$
$$\text{a 95\% ci for p would be : } \hat{p}_{MLE} \pm Z_{0.975}\sqrt{\frac{\hat{p}_{MLE}(1-\hat{p}_{MLE})}{25}} = 0.48 \pm 0.1958395 = (0.284, 0.676)$$

- we can see that the CI and CS are rather close

Q6

Question 6

The waiting time for a bus at a given corner at a certain time of day is known to have a uniform distribution $U[0,\theta]$. From other similar routes it is known that θ has a Pareto distribution Pa(7,4), where the density of Pareto distribution $Pa(\alpha,\beta)$ is $\pi_{\alpha,\beta}(\theta)=(\alpha/\beta)(\beta/\theta)^{\alpha+1}$ for $\theta \geq \beta$ and 0, otherwise. Waiting times of 10, 3, 2, 5 and 14 minutes have been observed at the given corner at the last 5 days.

1. Show that the Pareto distribution provides a conjugate prior for uniform data and find the posterior distribution of θ.
2. Estimate θ w.r.t. to the quadratic error (recall that $E(Pa(\alpha,\beta))=\alpha\beta/(\alpha-1)$, $\alpha>1$).
3. Find a 95% HPD for θ.
4. Test the hypothesis H_0: $0 \leq \theta \leq 15$ vs. H_1: $\theta > 15$ by choosing the most likely (a posteriori) hypothesis.

Q6.1

$$\pi(\theta|y) = \frac{f(\theta|y)\pi(\theta)}{f(y)} \propto L(\theta;y)\pi(\theta) = \Pi_{i=1}^{n}\frac{1}{\theta}\frac{\alpha}{\beta}\left(\frac{\beta}{\theta}\right)^{\alpha+1}\mathbb{I}\{\theta \geq y_{max}\}\mathbb{I}\{\theta \geq \beta\} \propto$$
$$\propto \frac{\alpha+n}{\tilde{\beta}}\left(\frac{\tilde{\beta}}{\theta}\right)^{\alpha+n+1}\mathbb{I}\{\theta \geq y_{max}\}\mathbb{I}\{\theta \geq \beta\} \Rightarrow \pi(\theta|y) \sim Pa(\alpha+n, max(y_{max}, \beta))$$

Q6.2

$$\hat{\theta} = E[\theta|y] = \frac{(\alpha+n)max(y_{max},\beta)}{(\alpha+n)-1} = \frac{12*14}{11} = 15.27$$

Q6.3

$$\text{notice that the density function is a decrecing function of } \theta \text{ thus the HPD has a form of } (\tilde{\beta}, C_1)$$
$$\text{we want a 95\% HPD so } C_1 = Pa(12,14)_{0.95} = 17.97$$
$$\text{thus the 95\% HPD is: } (14, 17.97)$$

Q6.4

$$\pi_0 = P(\theta < 15) = 1 - (\frac{14}{15})^{12} = 1 - 0.437 = 0.563 > 0.5 \Rightarrow \text{we will not reject } H_0$$