Lecturer: Prof. Dr. Mete SONER
Coordinator: Yilin WANG

ETH Zürich

Solution Series 9

Q1. Let $\alpha, \beta > 0$, the p.d.f. of a beta distribution with parameters α and β is

$$f(x|\alpha,\beta) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} & \text{for } 0 < x < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Suppose that $X_1, \cdots, X_n$ form a random sample from the Bernoulli distribution with parameter θ, which is unknown $(0 < \theta < 1)$. Suppose also that the prior distribution of θ is the beta distribution with parameters $\alpha > 0$ and $\beta > 0$. Show that the posterior distribution of θ given that $X_i = x_i (i = 1, \dots, n)$ is the beta distribution with parameters $\alpha + \sum_{i=1}^n x_i$ and $\beta + n - \sum_{i=1}^n x_i$.

In particular the family of beta distributions is a conjugate family of prior distributions for samples from a Bernoulli distribution. If the prior distribution of θ is a beta distribution, then the posterior distribution at each stage of sampling will also be a beta distribution, regardless of the observed values in the sample.

Solution:

First we calculate the joint p.f. of $X_1, \cdots, X_n, \theta$:

$$\begin{aligned} f_{X_1,\cdots,X_n,\theta}(x_1,\cdots,x_n,x) &= x^y(1-x)^{n-y}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1} \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha+y-1}(1-x)^{\beta+n-y-1} \end{aligned}$$

where $y = x_1 + \cdots + x_n$. So that the marginal p.f. of $X_1, \cdots, X_n$ at $(x_1, \cdots, x_n)$ is

$$\int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha+y-1}(1-x)^{\beta+n-y-1}dx = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma(\alpha+y)\Gamma(\beta+n-y)}{\Gamma(\alpha+\beta+n)}.$$

Thus the conditional p.d.f. of θ given $x_1, \cdots, x_n$ is

$$f_{\theta|x_1,\cdots,x_n}(x) = \frac{\Gamma(\alpha+\beta+n)}{\Gamma(\alpha+y)\Gamma(\beta+n-y)}x^{\alpha+y-1}(1-x)^{\beta+n-y-1},$$

which is the Beta distribution with parameters $\alpha + y$ and $\beta + n - y$.

Q2. Let $\xi(\theta)$ be defined as follows: for constants $\alpha > 0$ and $\beta > 0$:

$$\xi(\theta) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{-(\alpha+1)}e^{-\beta/\theta} & \text{for } \theta > 0, \\ 0 & \text{for } \theta \leq 0. \end{cases}$$

A distribution with this p.d.f. is called an *inverse gamma distribution*.

(a) Verify that $\xi(\theta)$ is actually a p.d.f..

(b) Consider the family of probability distributions that can be represented by a p.d.f. $\xi(\theta)$ having the given form for all possible pairs of constants $\alpha > 0$ and $\beta > 0$. Show that this family is a conjugate family of prior distributions for samples from a normal distribution with a known value of the mean μ and an unknown value of the variance θ.

Solution:

(a) The integral of ξ is

$$\begin{aligned} &\int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{-(\alpha+1)}e^{-\beta/\theta}d\theta \\ =&\frac{\beta^\alpha}{\Gamma(\alpha)}\int_\infty^0 -\frac{\beta}{y^2}\beta^{-(\alpha+1)}y^{\alpha+1}e^{-y}dy \\ =&\frac{1}{\Gamma(\alpha)}\int_0^\infty y^{\alpha-1}e^{-y}dy \\ =&1 \end{aligned}$$

(b) As we did before, let Θ be r.v. following the inverse gamma distribution with parameter α and β, let X be a random variable such that the conditional p.d.f. given $\Theta = \theta$ is

$$f_{X|\theta}(x) = \frac{1}{\sqrt{2\pi\theta}}e^{-(x-\mu)^2/2\theta}.$$

The conditional p.d.f. of Θ given $X = x$ is

$$\begin{aligned} f_{\Theta|x}(\theta) &= \frac{\xi(\theta)f_{X|\theta}(x)}{\int_0^\infty \xi(\theta)f_{X|\theta}(x)d\theta} \\ &= \frac{\xi(\theta)f_{X|\theta}(x)}{\frac{\beta^\alpha}{\Gamma(\alpha)\sqrt{2\pi}}\int_0^\infty \theta^{-(\alpha+1/2)-1}e^{-(\beta+(x-\mu)^2/2)/\theta}d\theta} \\ &= \frac{\frac{\beta^\alpha}{\Gamma(\alpha)\sqrt{2\pi}}\theta^{-(\alpha+1/2)-1}e^{-(\beta+(x-\mu)^2/2)/\theta}}{\frac{\beta^\alpha}{\Gamma(\alpha)\sqrt{2\pi}}\frac{\Gamma(\alpha+1/2)}{[\beta+(x-\mu)^2/2]^{\alpha+1/2}}} \\ &= \frac{[\beta']^{\alpha'}}{\Gamma(\alpha')}\theta^{-\alpha'-1}e^{-\beta'/\theta}. \end{aligned}$$

Where we set $\alpha' = \alpha + 1/2$ and $\beta' = \beta + (x-\mu)^2/2$. The conditional distribution of Θ given $X = x$ is an inverse Gamma distribution with parameter α' and β'. Thus the family of inverse Gamma distribution is a family of prior distributions for samples from a normal distribution with a known mean μ and an unknown value of the variance.

Q3. Suppose that the number of defects in a 1200-foot roll of magnetic recording tape has a Poisson distribution for which the value of the mean θ is unknown, and the prior distribution of θ is the gamma distribution with parameters $\alpha = 3$ and $\beta = 1$. When five rolls of this tape are selected at random and inspected, the numbers of defects found on the rolls are $2, 2, 6, 0$ and 3.

(a) What is the posterior distribution of θ?

(b) If the squared error loss function is used, what is the Bayes estimate of θ?

Solution:

(a) Recall that X follows the Poisson distribution with mean θ then

$$\mathbb{P}(X = k) = e^{-\theta}\frac{\theta^k}{k!}.$$

And θ has the p.d.f.

$$f_\theta(x) = \frac{x^{3-1}e^{-x}}{\Gamma(3)} = x^2e^{-x}/2.$$

The conditional p.d.f. of θ is obtained as before:

$$\begin{aligned} f_{\theta|2,2,6,0,3}(x) &= \frac{x^2e^{-x}/2 \cdot e^{-5x}x^{2+2+6+0+3}/(2!2!6!0!3!)}{\int_0^\infty s^2e^{-s}/2 \cdot e^{-5s}s^{2+2+6+0+3}/(2!2!6!0!3!)ds} \\ &= \frac{e^{-6x}x^{15}}{\int_0^\infty e^{-6s}s^{15}ds} \\ &= e^{-6x}x^{15}\frac{\Gamma(16)}{6^{16}} := \xi(x|2,2,6,0,3) \end{aligned}$$

Remark also that the posterior distribution of θ is the Gamma distribution with parameter $(\alpha = 16, \beta = 6)$.

(b) The squared error loss function is

$$L(\theta, a) = (\theta - a)^2.$$

The Bayes estimator of θ is the a which minimizes

$$\mathbb{E}[L(\theta, a)|\text{observation}] = \int L(x, a)\xi(x|\text{observation})dx,$$

where $\xi(\theta|\text{observation})$ is the posterior p.d.f. of θ given the observation.
We have computed the k-th moments of Gamma distribution X with parameter α and β in previous series, in particular

$$\mathbb{E}(X) = \frac{\alpha}{\beta}, \quad \mathbb{E}(X^2) = \frac{\alpha(\alpha+1)}{\beta^2}.$$

Hence

$$\begin{aligned}\mathbb{E}[(\theta-a)^2|2,2,6,0,3] &= \mathbb{E}[\theta^2|2,2,6,0,3] - 2a\mathbb{E}[\theta|2,2,6,0,3] + a^2\\ &= \frac{17\times 16}{6^2} - 2a\frac{16}{6} + a^2\\ &= \left(a-\frac{8}{3}\right)^2 + \frac{4}{9}.\end{aligned}$$

The Bayes estimate for the observation 2, 2, 6, 0, 3 is 8/3.

Q4. Let $c > 0$ and consider the loss function

$$L(\theta,a) = \begin{cases} c|\theta-a| & \text{if } \theta < a,\\ |\theta-a| & \text{if } \theta \geq a.\end{cases}$$

Assume that θ has a continuous distribution.

(a) Let $a \leq q$ be two real numbers, show that

$$\mathbb{E}[L(\theta,a)-L(\theta,q)] \geq (q-a)[\mathbb{P}(\theta\geq q) - c\mathbb{P}(\theta\leq q)],$$

and

$$\mathbb{E}[L(\theta,a)-L(\theta,q)] \leq (q-a)[\mathbb{P}(a\leq\theta) - c\mathbb{P}(\theta\leq a)].$$

(b) Prove that a Bayes estimator of θ will be any $1/(1+c)$ quantile of the posterior distribution of θ.

Solution: Let θ follows a continuous distribution,

(a) let $a \leq q$,

$$\begin{aligned}\mathbb{E}[L(\theta,a)-L(\theta,q)] =& \mathbb{E}[(L(\theta,a)-L(\theta,q))(\mathbf{1}_{\theta\leq a} + \mathbf{1}_{a\leq\theta\leq q} + \mathbf{1}_{q\leq\theta})]\\ =& \mathbb{E}[c(a-\theta-q+\theta)\mathbf{1}_{\theta\leq a}] + \mathbb{E}[(\theta-a-c(q-\theta))\mathbf{1}_{a\leq\theta\leq q}]\\ &+ \mathbb{E}[(\theta-a-\theta+q)\mathbf{1}_{q\leq\theta}]\\ =& c(a-q)\mathbb{P}(\theta\leq a) + (1+c)\mathbb{E}(\theta\mathbf{1}_{a\leq\theta\leq q}) - (a+cq)\mathbb{P}(a\leq\theta\leq q)\\ &+ (q-a)\mathbb{P}(q\leq\theta)\\ \geq& (q-a)[\mathbb{P}(q\leq\theta) - c\mathbb{P}(\theta\leq a)] + a(1+c)\mathbb{P}(a\leq\theta\leq q)\\ &- (a+cq)\mathbb{P}(a\leq\theta\leq q)\\ =& (q-a)[\mathbb{P}(q\leq\theta) - c\mathbb{P}(\theta\leq q)].\end{aligned}$$

Similarly we have

$$\mathbb{E}[L(\theta,a)-L(\theta,q)] \leq (q-a)[\mathbb{P}(a\leq\theta) - c\mathbb{P}(\theta\leq a)].$$

(b) Let q be a $1/(1+c)$-quantile of θ. Then

$$\mathbb{P}(\theta \leq q) = 1/(1+c) \text{ and } \mathbb{P}(\theta \geq q) = c/(1+c).$$

For all $a \leq q$, by the first inequality,

$$\mathbb{E}(L(\theta, a) - L(\theta, q)) \geq 0.$$

For all $a \geq q$, by the second inequality,

$$E(L(\theta, a) - L(\theta, q)) \geq (q - a)[\mathbb{P}(q \leq \theta) - c\mathbb{P}(\theta \leq q)] = 0.$$

Thus q minimizes $\mathbb{E}[L(\theta, a)]$ among all $a \in \mathbb{R}$, is the Bayes estimator with loss function L.