University of California, Santa Cruz
Department of Applied Mathematics and Statistics
Baskin School of Engineering
Classical and Bayesian Inference - AMS 132

Homework 2

Instructions: You have until Friday, February 2, to complete the assignment. It has to be returned during 10 last minutes of class (4:55 pm to 5:05 pm) or between 1:00 pm and 3:00 pm in office BE 357B.

This homework includes a BONUS exercise. The score in that exercise can replace the score of ANY other exercise you choose, in this or another homework .

1. 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 α and β. **[15 pts.]**

   (a) Show that the posterior mean is a weighted average of the prior mean and the sample mean, with weights denoted by γ_n, this means, the posterior mean can be written as $\gamma_n E(\theta) + (1-\gamma_n)\overline{X}_n$, where $\overline{X}_n = \sum_{i=1}^n X_i/n$ and show that $\gamma_n \to 0$ as the sample size increases, this is, as $n \to \infty$. **[3 pts.]**
   Since the gamma prior is conjugate for Poisson sampling, we have that $\theta \mid \boldsymbol{x} \sim Gamma(\alpha + \sum_{i=1}^n x_i, \beta + n)$, and the posterior mean is given by

   $$E(\theta \mid \boldsymbol{x}) = \frac{\alpha + \sum_{i=1}^n x_i}{\beta + n} = \frac{\beta}{\beta+n}\frac{\alpha}{\beta} + \left(1 - \frac{\beta}{\beta+n}\right)\overline{x}_n = \gamma_n E(\theta) + (1-\gamma_n)\overline{x}_n.$$

   Finally, $\lim_{n\to\infty} \gamma_n = \lim_{n\to\infty} \frac{\beta}{\beta+n} = 0$, this is, $\gamma_n \to 0$ as $n \to \infty$.

   (b) Considering the square error loss function, find Bayes estimator for θ, and show that they form a consistent sequence of estimators of θ. **[3 pts.]**
   We have to show that Bayes estimator converges in probability to θ as $n \to \infty$.
   By the law of large number we know that $\overline{X}_n$ converges in probability to $E(X) = \theta$, as $n \to \infty$.
   Now, under square error loss function Bayes estimator is the posterior mean, so, by (a) it follows that

   $$\delta^\star(\boldsymbol{X}) = \frac{\alpha + \sum_{i=1}^n X_i}{\beta + n} = \gamma_n E(\theta) + (1-\gamma_n)\overline{X}_n.$$

   Since $\gamma_n \to 0$ and $\overline{X}_n$ converges in probability to θ, as $n \to \infty$, it follows that Bayes estimator converges in probability to θ.

   (c) If the prior mean and prior variance are equal to 3, find the values of the parameters of the prior distribution. **[3 pts.]**
   Solving $E(\theta) = \alpha/\beta = 3$ and $Var(\theta) = \alpha\beta^2 = 3$, it follows that $\alpha = 3$ and $\beta = 1$.

(d) 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. Find Bayes estimate for θ under square error loss. **[3 pts.]**
We have that $\alpha = 3$, $\beta = 1$, $n = 5$, and $\sum_{i=1}^{5} x_i = 13$, therefore, by (b) we get that $\delta^\star(\boldsymbol{x}) = \frac{3+13}{1+5} = \frac{16}{6}$.

(e) Find Bayes estimate for the probability of a 1200-foot roll of magnetic recording tape having no defects. **[3 pts.]**
Note that what we want to estimate is $P(X = 0) = e^{-\theta}\theta^0/0! = e^{\theta}$. So we want to estimate parameter $\Psi = h(\theta) = e^{-\theta}$, which under square error loss is the posterior mean. So,

$$\delta^\star(\Psi) = E(e^{-\theta} \mid \boldsymbol{x}) = \int_0^\infty e^{-\theta}\xi(\theta \mid \boldsymbol{x})d\theta = \int_0^\infty e^{-\theta}\frac{(\beta+n)^{\alpha+\sum_{i=1}^n x_i}}{\Gamma(\alpha+\sum_{i=1}^n x_i)}\theta^{\alpha+\sum_{i=1}^n x_i - 1}e^{-(\beta+n)\theta}$$

$$= \left(\frac{\beta+n}{\beta+n+1}\right)^{\alpha+\sum_{i=1}^n x_i} = \frac{6}{7}^{14}.$$

2. Suppose that $X_1, \ldots, X_n$ form a random sample from an exponential distribution for which the value of the parameter θ is unknown. **[15 pts.]**

(a) Determine the maximum likelihood estimator and estimate of θ. **[5 pts.]**
The likelihood function is $f_n(\boldsymbol{x} \mod \theta) = \theta^n e^{-\theta\sum_{i=1}^n x_i}$. Its logarithm is $L(\theta) = n\log(\theta) - \theta\sum_{i=1}^n x_i$. Now we find the value of θ that maximizes $L(\theta)$:

$$\frac{d}{d\theta}L(\theta) = \frac{n}{\theta} - \sum_{i=1}^n x_i = 0 \Rightarrow \theta = \frac{n}{\sum_{i=1}^n x_i}.$$

Now we check that that value of θ is a maximum: $\frac{d^2}{d\theta^2}L(\theta) = -\frac{n}{\theta^2} < 0$, for all θ. Therefore, $\widehat{\theta} = \frac{n}{\sum_{i=1}^n X_i}$ is the maximum likelihood estimator and $\widehat{\theta} = \frac{n}{\sum_{i=1}^n x_i}$ is the maximum likelihood estimate.

(b) Determine the maximum likelihood estimator and estimate of the probability of observing a value equal or smaller than x_0, where $x_0 > 0$. **[5 pts.]**
Note that we need to estimate $g(\theta) = P(X \le x_0) = \int_0^{x_0}\theta e^{-\theta x}dx = -e^{-\theta x}|_0^{x_0} = 1 - e^{-\theta x_0}$. By the invariance property of the M.L.E. we get that the maximum likelihood estimator of $P(X \le x_0)$ is $g(\widehat{\theta}) = 1 - e^{-\widehat{\theta}x_0} = 1 - e^{-nx_0/\sum_{i=1}^n X_i}$ and the maximum likelihood estimate is $g(\widehat{\theta}) = 1 - e^{-nx_0/\sum_{i=1}^n x_i}$.

(c) Determine the maximum likelihood estimator and estimate of the median of the distribution. **[5 pts.]**
We know that the median, say m, for a continuous random variable is $P(X \le m) = 0.5$. Then, by (b), it follows that the median is found by solving $1 - e^{-\theta m} = 0.5$, therefore $m = -log(0.5)/\theta = log(2)/\theta$ and the maximum likelihood estimator of the median is $m = g(\widehat{\theta}) = \log(2)/\widehat{\theta} = \log(2)\sum_{i=1}^n X_i/n$. The maximum likelihood estimate of the median is $m = g(\widehat{\theta}) = \log(2)\sum_{i=1}^n x_i/n$

3. Suppose that $X_1, \ldots, X_n$ form a random sample from a gamma distribution with parameters a and θ, where a is known and θ is unknown. It is known that θ can only be 1, or 2, or 3, and $a = 3$. Six random variables are observed to be $0.5, 0.8, 1.2, 0.3, 1.4, 0.2$. **[15 pts.]**

   (a) Write the statistical model. **[4 pts.]**
   Let X be the random variable of interest. The p.d.f. is $f(x \mid \theta) = \frac{\theta^a}{\Gamma(a)} x^{a-1} e^{-\theta x}$, $\theta \in \Omega$, where $\Omega = \{1, 2, 3\}$.

   (b) Plot the likelihood function and plot the logarithm of the likelihood function. **[4 pts.]**
   We have that $n = 6$, $\sum_{i=1}^{6} x_i = 4.4$, $a = 3$, and $\prod_{i=1}^{n} x_i^{a-1} = 0.0016$. So the likelihood function and its logarithm are

$$f_n(\boldsymbol{x} \mid \theta) = \frac{\theta^{na}}{\Gamma(a)^n} \prod_{i=1}^{n} x_i^{a-1} e^{-\theta \sum_{i=1}^{n} x_i} = \frac{\theta^{18}}{64} \times 0.0016 e^{-4.4\theta},$$

$$\begin{aligned} L(\theta) &= na \log(\theta) - n \log(\Gamma(a)) + (a-1) \sum_{i=1}^{n} \log(x_i) - \theta \sum_{i=1}^{n} x_i, \\ &= 18 \log(\theta) - 6 \log(2) + 2(-3.2109) - 4.4\theta. \end{aligned}$$

[Figure 1 about here.]

   (c) Find the maximum likelihood estimator and estimate of θ. **[3 pts.]**
   From the plot is easy to see that the likelihood function (and its logarithm) is maximized when $\theta = 3$. So $\widehat{\theta} = 3$ is the M.L.E.

   (d) Find the maximum likelihood estimator and estimate of $2\theta + 8$. **[4 pts.]**
   Since we are interested in the M.L.E. of $g(\theta) = 2\theta + 8$it follows that $g(\widehat{\theta}) = 2 * 3 + 8 = 14$.

4. Suppose that $X_1, \ldots, X_n$ form a random sample from a gamma distribution with parameters a and θ, where a is known and θ is unknown. Suppose that a discrete prior for θ is assumed, where $\xi(1) = 0.3$, $\xi(2) = 0.5$, and $\xi(3) = 0.2$. Six random variables are observed to be $0.5, 0.8, 1.2, 0.3, 1.4, 0.2$. **[15 pts.]**

   (a) Find the posterior distribution of θ. **[4 pts.]**
   From the definition of posterior distribution and exercise 4 we have that

$$\xi(\theta \mid \boldsymbol{x}) = \frac{f_n(\boldsymbol{x} \mid \theta)\xi(\theta)}{\sum_{\theta \in \{1,2,3\}} f_n(\boldsymbol{x} \mid \theta)\xi(\theta)} = \frac{\frac{\theta^{18}}{64} \times 0.0016 e^{-4.4\theta} \xi(\theta)}{\sum_{\theta \in \{1,2,3\}} \frac{\theta^{18}}{64} \times 0.0016 e^{-4.4\theta} \xi(\theta)} = \frac{\frac{\theta^{18}}{64} \times 0.0016 e^{-4.4\theta} \xi(\theta)}{0.0041}.$$

   Therefore, the posterior distribution is given by

$$\xi(1 \mid \boldsymbol{x}) = \frac{\frac{1^{18}}{64} \times 0.0016 e^{-4.4*1} \times 0.3}{0.0041} = 0.0000225,$$

$$\xi(2 \mid \boldsymbol{x}) = \frac{\frac{2^{18}}{64} \times 0.0016 e^{-4.4*2} \times 0.5}{0.0041} = 0.1210,$$
$$\xi(3 \mid \boldsymbol{x}) = 1 - \xi(1 \mid \boldsymbol{x}) - \xi(1 \mid \boldsymbol{x}) = 0.8789.$$

(b) Find Bayes estimate under absolute error loss function. **[4 pts.]**
Under absolute error loss function Bayes estimate is the median. For discrete random variables, the median, say m, is given by $P(\theta \leq m) \geq 0.5$ and $P(\theta \geq m) \geq 0.5$. From the posterior distribution of θ given in (a), we have that Bayes estimate is $\delta^\star(\boldsymbol{x}) = 3$.

(c) Find Bayes estimate under square error loss function. **[4 pts.]**
Under square error loss function Bayes estimate is the posterior mean, therefore

$$\delta^\star(\boldsymbol{x}) = E(\theta \mid \boldsymbol{x}) = \sum_{\theta \in \{1,2,\}} \theta\xi(\theta \mid \boldsymbol{x}) = 1 * 0.0000225 + 2 * 0.1210 + 3 * 0.8789 = 2.878723.$$

(d) Find Bayes estimate for $\Psi = 2\theta + 8$ under square error loss function. **[3 pts.]**
Under square error loss function Bayes estimate is the posterior mean, therefore

$$\delta^\star(\boldsymbol{x}) = E(\Psi \mid \boldsymbol{x}) = E(2\theta + 8 \mid \boldsymbol{x}) = 2E(\theta \mid \boldsymbol{x}) + 8 = 13.75745.$$

5. (**Bonus exercise**): Find the distribution of a new observation given and observed sample of size n, $X_1 = x_1, \ldots, X_n = x_n$, for the following models: **[15 pts.]**

- $X_1, \ldots, X_n$ form a random sample from the Bernoulli distribution with parameter θ, and θ has a beta prior distribution with parameters $\alpha > 0$ and $\beta > 0$. **[4 pts.]**
Since the beta prior is conjugate for Bernoulli sampling, it follows that

$$\theta \mid \boldsymbol{x} \sim Beta(\alpha + \sum_{i=1}^n x_i, \beta + n - \sum_{i=1}^n x_i).$$

The p.d.f. of a new observation given $\boldsymbol{x}$ is given by

$$f(x_{n+1} \mid \boldsymbol{x}) = \int_0^1 \theta^{x_{n+1}}(1-\theta)^{1-x_{n+1}}\xi(\theta \mid \boldsymbol{x})d\theta,$$

$$= \frac{\Gamma(\alpha + \sum_{i=1}^n x_i + x_{n+1})\Gamma(\beta + n - \sum_{i=1}^n x_i - x_{n+1} + 1)\Gamma(\alpha + \beta + n)}{\Gamma(\alpha + \sum_{i=1}^n x_i)\Gamma(\beta + n - \sum_{i=1}^n x_i)\Gamma(\alpha + \beta + n + 1)}.$$

- $X_1, \ldots, X_n$ form a random sample from the Poisson distribution with parameter θ, and θ has a gamma prior distribution with parameters $\alpha > 0$ and $\beta > 0$. **[4 pts.]**
Since the gamma prior is conjugate for Poisson sampling, it follows that

$$\theta \mid \boldsymbol{x} \sim Gamma(\alpha + \sum_{i=1}^n x_i, \beta + n).$$

The p.d.f. of a new observation given $\boldsymbol{x}$ is given by

$$f(x_{n+1} \mid \boldsymbol{x}) = \int_0^\infty \frac{e^{-\theta}\theta^{x_{n+1}}}{x_{n+1}!}\xi(\theta \mid \boldsymbol{x})d\theta,$$

$$= \frac{(\beta+n)^{\alpha + \sum_{i=1}^n x_i}\Gamma(\alpha + \sum_{i=1}^n x_i + x_{n+1})}{(\beta + n + 1)^{\alpha + \sum_{i=1}^n x_i + x_{n+1}}\Gamma(\alpha + \sum_{i=1}^n x_i)x_{n+1}!}.$$

- $X_1, \ldots, X_n$ form a random sample from the Normal distribution with unknown mean θ and known variance $\sigma^2 > 0$, and θ has a normal prior distribution with mean μ_0 and variance v_0^2. **[3 pts.]**

  Since the normal prior is conjugate for normal sampling with known variance, it follows that

$$\theta \mid \boldsymbol{x} \sim Normal(\mu_1, v_1^2),$$

  where $\mu_1 = \frac{\sigma^2\mu_0 + nv_0^2\overline{x}_n}{\sigma^2 + nv_0^2}$, and $v_1^2 = \frac{\sigma^2 v_0^2}{\sigma^2 + nv_0^2}$. The p.d.f. of a new observation given $\boldsymbol{x}$ is given by

$$f(x_{n+1} \mid \boldsymbol{x}) = \int_0^\infty \frac{1}{\sqrt{2\pi v_1^2}} \exp\left\{-\frac{1}{2v_1^2}(\theta - \mu_1)^2\right\} \xi(\theta \mid \boldsymbol{x}) d\theta,$$

$$= \frac{1}{\sqrt{2\pi(\sigma^2 + v_1^2)}} \exp\left\{-\frac{1}{2(\sigma^2 + v_1^2)}(\theta - \mu_1)^2\right\}.$$

- $X_1, \ldots, X_n$ form a random sample from the exponential distribution with parameter θ, and θ has a gamma prior distribution with parameters $\alpha > 0$ and $\beta > 0$. **[4 pts.]**

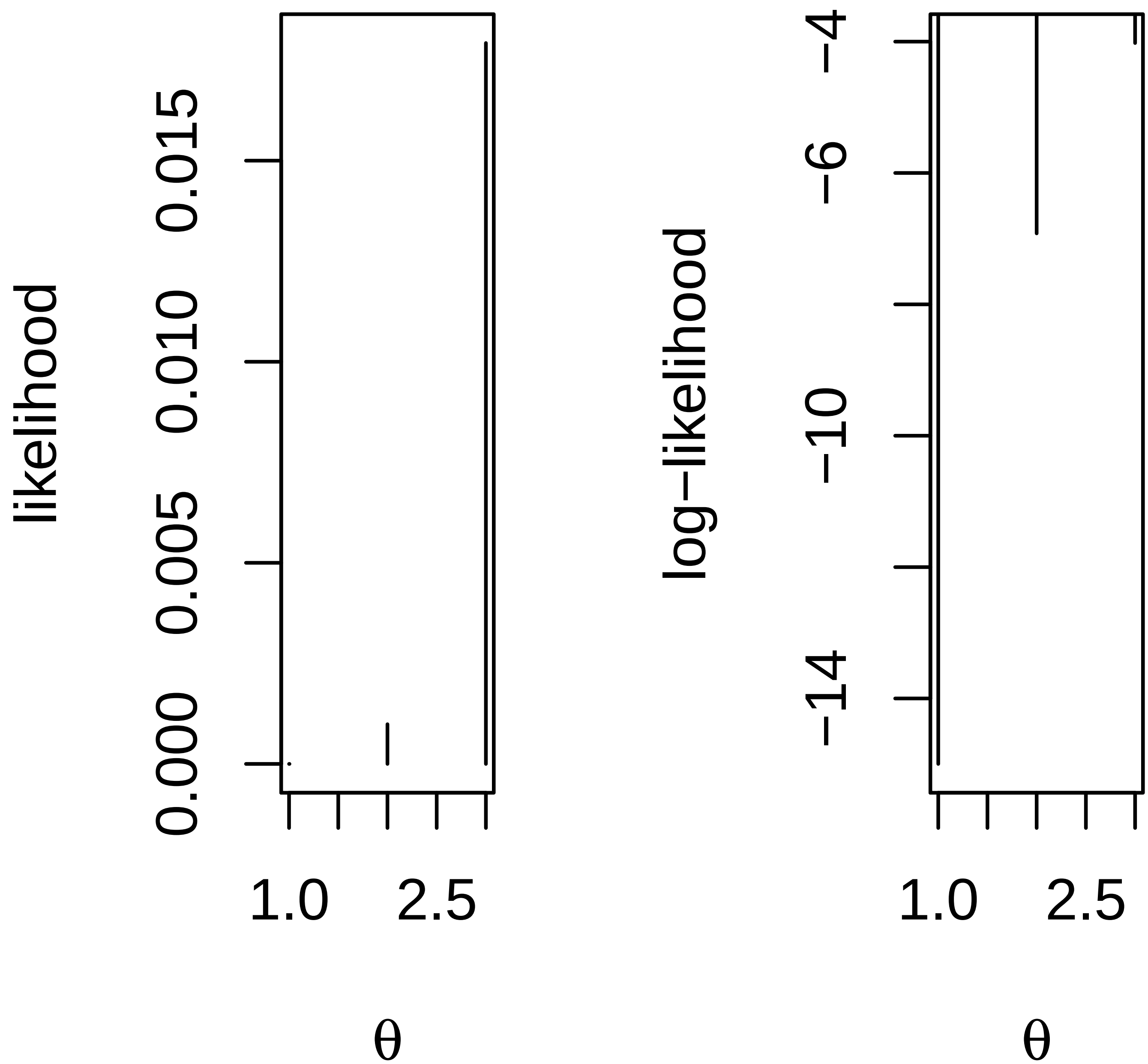
  Since the gamma prior is conjugate for exponential sampling, it follows that

$$\theta \mid \boldsymbol{x} \sim Gamma(\alpha + n, \beta + \sum_{i=1}^n x_i).$$

  The p.d.f. of a new observation given $\boldsymbol{x}$ is given by

$$f(x_{n+1} \mid \boldsymbol{x}) = \int_0^\infty \theta e^{-\theta x_{n+1}} \xi(\theta \mid \boldsymbol{x}) d\theta,$$

$$= \frac{(\alpha + n)(\beta + \sum_{i=1}^n x_i)^{\alpha+n}}{(\beta + \sum_{i=1}^n x_i + x_{n+1})^{\alpha+n+1}}.$$


likelihood
0.000
0.005
0.010
0.015
1.0
2.5
θ
log−likelihood
−14
−10
−6
−4
1.0
2.5
θ