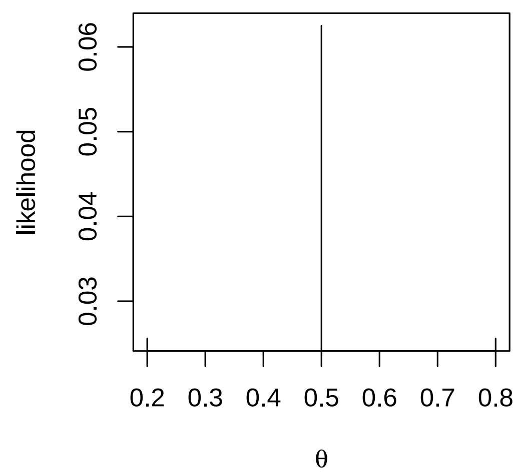
University of California, Santa Cruz
Department of Applied Mathematics and Statistics
Baskin School of Engineering
Classical and Bayesian Inference - AMS 132

Exam 1: Class, Version A

Name:

1. Suppose that $X_1, \ldots, X_n$ are random variables describing whether a student approves its AMS 132 class or not. Assume that $X_1, \ldots, X_n$ form a random sample from the Bernoulli distribution with unknown probability θ, where θ can be $0.2, 0.5$, or 0.8. A small random sample of 4 students was observed and 2 of them approved the class. **[15 pts]**

   (a) Plot the likelihood function. Specify the axes labels and values. **[5 pts]**



   (b) Find the maximum likelihood estimate of θ. **[5 pts]**
   The maximum likelihood estimate is $\widehat{\theta} = 0.5$

   (c) Find the maximum likelihood estimate of the variance of X_1. **[5 pts]**
   Notice that $Var(X_1) = \theta(1-\theta) = g(\theta)$. Then, by the invariance property of the MLE we have that the maximum likelihood estimate of $Var(X_1)$ is $g(\widehat{\theta}) = 0.5(1-0.5) = 0.25$

2. Suppose that $X_1, \ldots, X_8$ are random variables describing the monthly amount of accumulated rainfall during winter in Santa Cruz. Assume that $X_1, \ldots, X_8$ form a random sample from a gamma distribution with parameters $a > 0$ and $\theta > 0$, where a is known and θ is unknown. the density function is given by

$$f(x \mid a, \theta) = \frac{\theta^a}{\Gamma(a)} x^{a-1} e^{-\theta x}, \ \ x > 0.$$

Consider $a = 3$ and that the total amount of rainfall during winter is 10, this is, $\sum_{i=1}^{8} x_i = 10$. **[15 pts]**

(a) Write the statistical model. **[5 pts]**
Let X be the random variable describing the monthly amount of accumulated rainfall during winter in Santa Cruz. The joint distribution of the random variables is

$$f(\boldsymbol{x} \mid \theta) = \prod_{i=1}^{n} \frac{\theta^a}{\Gamma(a)} x^{a-1} e^{-\theta x} = \frac{\theta^{na}}{\Gamma(a)^n} \prod_{i=1}^{n} x_i^{a-1} e^{-\theta \sum_{i=1}^{n} x_i},$$

where $\theta \in \Omega = \mathbb{R}^+$.

(b) Consider an exponential prior distribution with mean equal to $1/2$ for θ. Considering square error loss function, find the Bayes estimate for θ. **[5 pts]**
Assuming that $\theta \sim exp(\alpha)$, and noting that $E(\theta) = 1/2$, it follows that $\theta \sim exp(2)$. First, lets find the posterior distribution of θ

$$\xi(\theta \mid \boldsymbol{x}) \propto \prod_{i=1}^{n} \frac{\theta^a}{\Gamma(a)} x_i^{a-1} e^{-\theta x_i} \xi(\theta) \propto \theta^{na} e^{-\theta \sum_{i=1}^{n} x_i} e^{-2\theta} \propto \theta^{na} e^{-\theta(2+\sum_{i=1}^{n} x_i)}.$$

Therefore, $\theta \mid \boldsymbol{x} \sim Gamma(na+1, 2+\sum_{i=1}^{n} x_i)$. Considering square error loss, Bayes estimate is the posterior mean. So, Bayes estimate is given by

$$\delta^\star = \frac{na+1}{2+\sum_{i=1}^{n} x_i} = \frac{31}{12}.$$

(c) Under the same prior distribution for θ as in (b), and considering square error loss function, find the Bayes estimate for the variance of the monthly amount of accumulated rainfall during winter in Santa Cruz. **[5 pts]**
The variance of the monthly amount of accumulated rainfall during winter in Santa Cruz is given by $Var(X) = \frac{a}{\theta^2} = \frac{3}{\theta^2}$. Here we want to estimate parameter $\Psi = \frac{3}{\theta^2}$. Considering square error loss, Bayes estimate is the posterior mean of Ψ. So,

$$\delta^\star = E(\frac{3}{\theta^2} \mid \boldsymbol{x}) = \int_0^\infty \frac{3}{\theta^2} \frac{12^{31}}{\Gamma(31)} \theta^{30} e^{-12\theta} d\theta = \frac{3 * 12^{31}}{\Gamma(31)} \int_0^\infty \theta^{28} e^{-12\theta} d\theta,$$

$$= \frac{3 * 12^{31}}{\Gamma(31)} \frac{\Gamma(29)}{12^{29}} = \frac{3 * 12^2}{29 * 30}.$$

Exam 1: Take home, Version A

Due date is Wednesday, February 7, at 4:00 pm, at Merrill Acad 102 (before the class starts). Late take home part will not be accepted!

1. Suppose that $X_1, \dots, X_n$ form a random sample from a Poisson distribution with unknown parameter θ. **[15 pts]**

   (a) Write $E(X_1^2)$ as a function of the parameter θ. **[3 pts]**
   Hint: Recall that $Var(X) = E(X^2) - [E(X)]^2$.
   Since $E(X) = Var(X) = \theta$, X_i for a random sample, and $E(X^2) = Var(X) + [E(X)]^2$, it follows that $E(X_1^2) = E(X^2) = \theta + \theta^2 = \theta(1 + \theta)$.

   (b) Find the maximum likelihood estimate of θ. **[4 pts]**
   The likelihood function of θ is

   $$f_n(\boldsymbol{x} \mid \theta) = \frac{e^{-n\theta}\theta^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}.$$

   Its logarithm is

   $$L(\theta) = -n\theta + \sum_{i=1}^n x_i \log(\theta) - \log(1/\prod_{i=1}^n x_i!).$$

   It derivative is

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

   Making the derivative equal to 0 and solving the equation we get that a candidate for being the maximum likelihood estimate of θ is

   $$\widehat{\theta} = \overline{X}_n = \frac{1}{n}\sum_{i=1}^n x_i.$$

   Finally, checking that

   $$\frac{d^2}{d\theta^2}L(\theta) = -\frac{1}{\theta^2}\sum_{i=1}^n x_i < 0,$$

   for every sample and θ, we conclude that $\widehat{\theta}$ is the maximum likelihood estimate of θ.

   (c) Find the maximum likelihood estimate of $E(X_1^2)$. **[4 pts]**
   From part (a) it follows that $E(X_1^2) = \theta(1+\theta) = g(\theta)$. By the invariance property of the M.L.E., it follows that the maximu likelihood estimate of $E(X_1^2)$ is $g(\widehat{\theta}) = \widehat{\theta}(1 + \widehat{\theta}) = \overline{X}_n(1 + \overline{X}_n)$.

   (d) Assume now that θ is a random parameter and has a gamma prior distribution with parameters $a > 0$ and $b > 0$. Consider square error loss function and find the Bayes estimate for θ. Under what kind of prior distribution are the maximum likelihood estimate and the Bayes estimate equals? **[4 pts]**

Since θ has a Gamma prior distribution and X_i follows a Poisson distribution, by theorem from class, it follows that

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

Under square error loss function, Bayes estimate, $\delta^\star$, is the posterior mean, therefore

$$\delta^\star = \frac{a + \sum_{i=1}^n x_i}{b + n}.$$

Finally, if $a = b = 0$, Bayes estimate and the maximum likelihood estimate are equal. This is, Bayes estimate and the maximum likelihood estimate are equal when an improper prior distribution for θ is considered.

2. Suppose that $X_1, \ldots, X_n$ form a random sample from a distribution that has p.d.f. given by

$$f(x \mid \theta) = \theta x e^{-0.5\theta x^2}, \;\; x > 0, \;\; \theta > 0.$$

Assume that θ is a random parameter with gamma prior distribution with parameters $a > 0$ and $b > 0$. **[15 pts]**

(a) Show that the gamma prior distribution is a conjugate prior for samples from the distribution that has p.d.f. as given above. Explain with your <u>own words</u> what this means. **[7 pts]**

Note that

$$\xi(\theta \mid \boldsymbol{x}) \propto \theta^n e^{-0.5\sum_{i=1}^n x_i^2} \theta^{a-1} e^{-b\theta} \propto \theta^{a+n-1} e^{-(0.5\sum_{i=1}^n x_i^2 + b)\theta},$$

therefore, $\theta \mid \boldsymbol{x} \sim Gamma(a + n, 0.5\sum_{i=1}^n x_i^2 + b)$.

This means, that if the parameter has a gamma prior distribution and the random variables have p.d.f. has defined above, then the posterior distribution of θ has also a gamma distribution.

(b) Find the p.d.f. of a new observation, X_{n+1}, given that $X_1 = x_1, \ldots, X_n = x_n$ have been observed, this is, find $f(x_{n+1} \mid x_1, \ldots, x_n)$. **[8 pts]**

$$\begin{aligned}
f(x_{n+1} \mid \boldsymbol{x}) &= \int_0^\infty \theta x_{n+1} e^{-0.5\theta x_{n+1}^2} \frac{(0.5\sum_{i=1}^n x_i^2 + b)^{a+n}}{\Gamma(a+n)} \theta^{a+n-1} e^{-(0.5\sum_{i=1}^n x_i^2 + b)\theta} d\theta, \\
&= x_{n+1} \frac{(0.5\sum_{i=1}^n x_i^2 + b)^{a+n}}{\Gamma(a+n)} \int_0^\infty \theta^{a+n+1-1} e^{-(0.5[\sum_{i=1}^n x_i^2 + x_{n+1}^2] + b)\theta} d\theta, \\
&= x_{n+1} \frac{(0.5\sum_{i=1}^n x_i^2 + b)^{a+n}}{\Gamma(a+n)} \frac{\Gamma(a+n+1)}{(0.5[\sum_{i=1}^n x_i^2 + x_{n+1}^2] + b)^{a+n+1}}, \\
&= x_{n+1}(a+n) \frac{(0.5\sum_{i=1}^n x_i^2 + b)^{a+n}}{(0.5[\sum_{i=1}^n x_i^2 + x_{n+1}^2] + b)^{a+n+1}},
\end{aligned}$$

$x_{n+1} > 0$.