

MAS3317

NEWCASTLE UNIVERSITY

SCHOOL OF MATHEMATICS & STATISTICS

SEMESTER 2 2013/14

MAS3317

Introduction to Bayesian Statistics

Time allowed: 1 hour 30 minutes

Candidates should attempt all questions. Marks for each question are indicated. However you are advised that marks indicate the relative weight of individual questions, they do not correspond directly to marks on the University scale.

There are FOUR questions in Section A and THREE questions in Section B.

Answers to questions should be entered directly on this question paper in the spaces provided. Rough work should be done on the blank sides of the pages. The rough work will not be marked. This question paper must be handed in, attached inside an anonymised cover sheet, at the end of the examination.

Beta distribution: If $X \sim Be(\alpha, \beta)$ then it has density

$$f(x|\alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}, \quad 0 < x < 1, \quad \alpha > 0, \beta > 0.$$

Also, $E(X) = \alpha/(\alpha + \beta)$, $Var(X) = \alpha\beta/\{(\alpha + \beta)^2(\alpha + \beta + 1)\}$ and $Mode(X) = (\alpha - 1)/(\alpha + \beta - 2)$.

Beta-Binomial distribution: If $X \sim BetaBin(n, \alpha, \beta)$ then it has density

$$f(x|n, \alpha, \beta) = {}^nC_x \frac{B(x + \alpha, \beta + n - x)}{B(\alpha, \beta)}, \quad x = 0, \dots, n, \quad \alpha > 0, \beta > 0.$$

Also, $E(X) = n\alpha/(\alpha + \beta)$, $Var(X) = n\alpha\beta(\alpha + \beta + n)(\alpha + \beta)^2(\alpha + \beta + 1)$.

Exponential distribution: If $X \sim Exp(\lambda)$ then it has density

$$f(x|\lambda) = \lambda e^{-\lambda x}, \quad x > 0, \quad \lambda > 0.$$

Also, $E(X) = 1/\lambda$ and $Var(X) = 1/\lambda^2$.

Gamma distribution: If $X \sim Ga(\alpha, \lambda)$ then it has density

$$f(x|\alpha, \lambda) = \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}, \quad x > 0, \quad \alpha > 0, \lambda > 0.$$

Also, $E(X) = \alpha/\lambda$ and $Var(X) = \alpha/\lambda^2$.

Normal distribution: If $X \sim N(\mu, 1/\tau)$ then it has density

$$f(x|\mu, \tau) = \left(\frac{\tau}{2\pi}\right)^{1/2} \exp\left\{-\frac{\tau}{2}(x - \mu)^2\right\}, \quad -\infty < x < \infty, \quad -\infty < \mu < \infty, \tau > 0.$$

Also, $E(X) = \mu$ and $Var(X) = 1/\tau$.

Note: If $X \sim N(0, 1)$ then the following quantiles apply:

x	1.2816	1.6449	1.9600	2.3263	2.5758
$\Pr(X < x)$	0.9	0.95	0.975	0.99	0.995

Poisson distribution: If $X \sim Po(\lambda)$ then it has probability function

$$f(x|\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, \dots, \quad \lambda > 0.$$

Also, $E(X) = \lambda$ and $Var(X) = \lambda$.

SECTION A

A1. A sports scientist measures the rate of oxygen consumption of some randomly chosen athletes immediately after exercise. She finds the following 95% frequentist confidence interval C_F for the population mean rate of oxygen consumption μ :

(1.997, 2.503) litres per minute.

In her report, she makes the following statement:

“My study revealed that the population mean rate of oxygen consumption for athletes lies in the range (1.997, 2.503) with probability 0.95”

(a) Briefly explain why the sports scientist’s interpretation of C_F is incorrect.

Answer:

(b) After formulating an appropriate prior distribution for μ , a Bayesian statistician updates his beliefs about μ using the data in the sports scientist’s study, giving

$$\mu|\mathbf{x} \sim N(2.39, 0.01).$$

(i) Find C_B , a 95% Bayesian confidence interval for μ .

Answer:

(ii) Briefly explain how C_B differs to C_F in interpretation.

Answer:

[Total A1: 9 marks]

A2. *Sapelo Island* is a barrier island off the southeast coast of the USA, and is on the migratory path of the rare black pelican. The random variable X represents the number of birds that are seen, per day, over the course of a week at an observation spot on the island.

(a) In the first week of April 2014, the following observations were made on X :

2 1 3 0 1 2 3

(i) Why might the Poisson distribution be a suitable model for X ?

Answer:

(ii) Assuming that $X \sim Po(\theta)$, obtain the likelihood function $f(\mathbf{x}|\theta)$ assuming your observations are independent.

Answer:

(b) After a consultation with an ornithologist (a zoologist with an expertise in birds), we are able to elicit a prior distribution for θ . Combining this prior distribution with the likelihood in part (a) results in the following posterior:

$$\theta|\mathbf{x} \sim Ga(18, 9).$$

Obtain the ornithologist's prior distribution for θ . How do the ornithologist's beliefs compare to our observations?

Answer:

[Part of this page has been left blank for your solutions to the last question]

- (c) Find α , the weight in the Bayes' linear rule for the posterior mean. Does this give more weight to the prior or the data?

Answer:

- (d) In a new analysis of the same data, a statistician attempts to use $\theta \sim N(3, 1.5)$. Briefly explain the similarities between this prior and the prior you obtained in part (b). What is the main problem with this new prior?

Answer:

[Total A4: 14 marks]

A3. Suppose we have a random sample from an exponential distribution, that is,
 $X_i|\theta \sim \text{Exp}(\theta)$, $i = 1, \dots, n$.

- (a) Use the Factorisation Theorem to show that $T = \sum_{i=1}^n X_i$ is sufficient for θ , and give an intuitive interpretation of T .

Answer:

- (b) Suppose we determine that a unit exponential distribution adequately represents our prior beliefs about the parameter θ . It can be shown that, after observing data $\mathbf{x}$, the posterior distribution for θ is

$$\theta|\mathbf{x} \sim \text{Ga}\left(n+1, \sum_{i=1}^n x_i + 1\right).$$

Use the fact that $T|\theta \sim \text{Ga}(n, \theta)$ to show that the same posterior is obtained knowing just the sufficient statistic T .

Answer:

[This page has been left blank for your solutions to the last question]

[Total A3: 11 marks]

- A4. (a) In the space below, briefly explain the terms *substantial prior knowledge* and *vague prior knowledge*. You may use diagrams to aid your descriptions.

SUBSTANTIAL PRIOR KNOWLEDGE

Answer:

VAGUE PRIOR KNOWLEDGE

Answer:

- (b) A *Jeffreys prior* is often used when we are *prior ignorant*. Give one advantage, and one disadvantage, of using such a prior.

Answer:

[Total A4: 6 marks]

SECTION B

B5. You work as a statistician for the Canadian spreadbetting company *ZenTrader*. You are interested in θ , the probability that the *Montreal Canadiens* ice hockey team win their next match against the *Boston Bruins*.

- (a) A discussion with a fan of the *Montreal Canadiens* results in the screenshot overleaf, taken from the *MATCH Uncertainty Elicitation Tool*. Briefly explain the benefits of the elicitation procedure used over the *bisection (quartile)* method, and write down the prior distribution for θ as suggested by *MATCH* (round your parameters to the nearest integer).

Answer:

- (b) The *Montreal Canadiens* have won two of their last six matches against the *Boston Bruins*. Given this information, find the posterior distribution for θ using the elicited prior from part (a).

Answer:



- (c) In the absence of any expert prior information, a *Haldane prior* for θ , with density

$$\pi(\theta) \propto \frac{1}{\theta(1-\theta)},$$

can be used.

- (i) Find the posterior distribution for θ using the Haldane prior.

Answer:

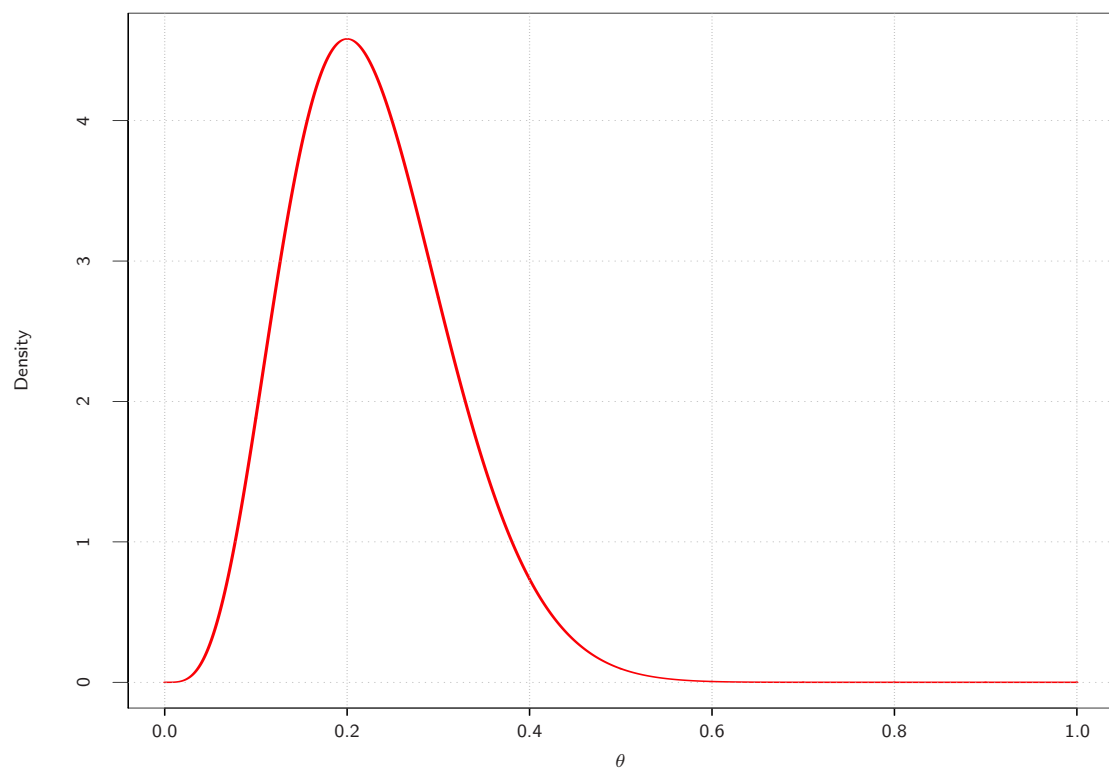
- (ii) Why might a Haldane prior be attractive in the absence of substantial prior information?

Answer:

- (d) Compare and contrast the posterior distributions for θ obtained in parts (b) and (c). You should use numerical summaries, as well as sketches of the posteriors using the graph paper overleaf, in your answer.

[Hint: the posterior resulting from the elicited prior has been drawn for you.]

Answer:



(d) Look at the following R code:

```
> g=function(a,b){
+   f=function(x){
+     l=x[1]
+     u=x[2]
+     (pbeta(u,a,b)-pbeta(l,a,b)-0.95)^2
+     +0.0001*(dbeta(u,a,b)-dbeta(l,a,b))^2
+   }
+   initial.l=qbeta(0.025,a,b)
+   initial.u=qbeta(0.975,a,b)
+   A=optim(c(initial.l,initial.u),f)
+   return(A$par)

> g(2,4)
[1] 0.02602514 0.67014432
> g(5,17)
[1] 0.06921317 0.39948449
```

Very briefly, explain what the function `g` does, and interpret the output for `g(2,4)` and `g(5,17)`. What does this tell us about the likely outcome of the next hockey match between the *Canadiens* and the *Bruins*?

Answer:

[Total B5: 20 marks]

B6. A random sample of size n is taken from a Normal $N(\mu, 1/\tau)$ distribution (where the precision τ is known) giving a sample mean $\bar{x}$. It can be shown that the likelihood function for μ is

$$f(\mathbf{x}|\mu) = \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left\{-\frac{n\tau}{2} [s^2 + (\bar{x} - \mu)^2]\right\},$$

where

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2.$$

- (a) Suppose our prior beliefs about μ can be summarised by a Normal $N(b, 1/d)$ distribution. Show that the posterior distribution is $\mu|\mathbf{x} \sim N(B, 1/D)$, where

$$B = \frac{db + n\tau\bar{x}}{D} \quad \text{and} \quad D = d + n\tau.$$

Answer:

[Part of this page has been left blank for your solution to the last question]

- (b) A sports scientist measures the rate of oxygen consumption of 10 randomly chosen athletes immediately after exercise. The sample mean is $\bar{x} = 2.25$ litres per minute. She assumes a Normal distribution for these measurements with a standard deviation of $1/\sqrt{6}$ litres per minute.

A prior study suggests that $\mu \sim N(2.6, 0.025)$.

- (i) Using this information, confirm the posterior distribution used in question A1(b), that is

$$\mu|\mathbf{x} \sim N(2.39, 0.01).$$

Answer:

- (ii) Find the posterior probability that the mean rate of oxygen consumption is greater than 2.518 litres per minute.
- (c) Obtain the posterior distribution for μ under the assumption of *vague prior knowledge*, and explain why the resulting 95% Bayesian confidence interval would be the same (numerically) as the sports scientist's frequentist confidence interval in question **A1**.

Answer:

(d) Given that

$$\frac{\partial}{\partial \mu} \log f(\mathbf{x}|\mu) = n\tau(\bar{x} - \mu),$$

show that, no matter what prior distribution is used for μ , the posterior is the same as that you obtained in part (c) *when n is large*.

Answer:

[Total B6: 20 marks]

- B7.** Suppose that X is the number of tropical storms occurring over the Atlantic Ocean in July 2014, and let θ be the daily rate of occurrence of such storms. A meteorologist assumes that $X|\theta \sim \text{Po}(31\theta)$, with the following prior distribution for θ :

Level of activity	θ	$\pi(\theta)$
Busy	0.25	0.2
Normal	0.1	0.8

- (a) Suppose that six tropical storms are observed over the Atlantic Ocean in July 2014. Determine the posterior distribution for θ .

Answer:

- (b) Suppose now that we want to predict the number of tropical storms Y over the Atlantic Ocean during August 2014. Determine the predictive probability function for Y , $f(y|x = 6)$, clearly stating any assumptions that you make.

Answer:

- (c) Complete the table below showing the posterior predictive distribution for Y , and use this table to find an approximate 90% prediction interval for Y .

Answer:

y	0	1	2	3	4	5	6
$f(y x = 6)$	0.0286	0.0894			0.1333	0.1049	0.0828
y	7	8	9	10	11	12	13
$f(y x = 6)$	0.0684	0.0573	0.0462		0.0243	0.0156	0.0093

- (d) Let ψ be the probability that a tropical storm develops into a hurricane. After a discussion with the meteorologist, we are able to elicit the following prior distribution for ψ :

$$\psi \sim Be(2, 23), \quad 0 < \psi < 1.$$

One of the last five tropical storms over the Atlantic Ocean developed into a hurricane. For the next five tropical storms observed in the future, show that the predictive probability distribution for the number that will develop into a hurricane is Beta-Binomial $BetaBin(5, 3, 27)$.

Answer:

- (e) Use Bayes' Theorem to show that the predictive probability distribution for future data y , given past data $\mathbf{x}$, can be written as

$$f(y|\mathbf{x}) = \frac{f(y|\psi)\pi(\psi|\mathbf{x})}{\pi(\psi|\mathbf{x}, y)},$$

assuming $\mathbf{x}$ and y are independent.

[Total B7: 20 marks]

THE END