

MAS2317/3317: Specimen exam paper

Dr. Lee Fawcett

Semester 2, 2013—2014

Revision material

Specimen exam paper: solutions

FREQUENTIST

■ *Description*

The probability of an outcome is the **relative frequency** with which the outcome would be obtained if the experiment were repeated a large number of times under similar conditions.

■ *Drawback*

Not obvious how many times the experiment should be repeated.

CLASSICAL

■ *Description*

This is based on the concept of **equally likely outcomes** resulting from ideas of symmetry.

If the outcome of an experiment must be one of n different outcomes and these n outcomes are equally likely, then the probability of each outcome is $1/n$.

■ *Drawback*

Only applies to equally likely outcomes.

SUBJECTIVE

■ *Description*

Your subjective probability for an outcome A represents your own judgement of the likelihood that the outcome will occur.

This judgement will be based on the beliefs and information you have at the time.

■ *Drawback*

It is not objective – but perhaps it is more obvious (honest) about when subjective beliefs are used.

Question A2(a)

If $E_1, E_2, \dots, E_n$ are a **partition** of $\mathcal{S}$ and F is any event with $\Pr(F) > 0$ then

$$\Pr(E_i|F) = \frac{\Pr(F|E_i)\Pr(E_i)}{\sum_{j=1}^n \Pr(F|E_j)\Pr(E_j)}, \quad i = 1, 2, \dots, n.$$

Question A2(a)

Using the definition of conditional probability, for $i = 1, 2, \dots, n$,

$$\Pr(E_i|F) = \frac{\Pr(E_i \cap F)}{\Pr(F)}$$

$$= \frac{\Pr(F \cap E_i)}{\Pr(F)}$$

$$= \frac{\Pr(F|E_i)\Pr(E_i)}{\Pr(F)} \quad \text{again using the definition}$$

$$= \frac{\Pr(F|E_i)\Pr(E_i)}{\sum_{j=1}^n \Pr(F|E_j)\Pr(E_j)} \quad \text{using the Law of Total Probability.}$$

Question A2(b)

Let

H = Does homework and P = Pass.

Then

$$\begin{aligned}\Pr(H|P) &= \frac{\Pr(P|H)\Pr(H)}{\Pr(P|H)\Pr(H) + \Pr(P|H^C)\Pr(H^C)} \\ &= \frac{0.9 \times 0.75}{0.9 \times 0.75 + 0.25 \times 0.25} \\ &= 0.915.\end{aligned}$$

Question 3(a), (b)

$$f(\mathbf{x}|\lambda) = \prod_{i=1}^n \frac{e^{-\lambda} \lambda^{x_i}}{x_i!} = \frac{e^{-n\lambda} \lambda^{\sum x_i}}{\prod x_i!}.$$

Also

$$\begin{aligned}\log f(\mathbf{x}|\lambda) &= -n\lambda + \sum x_i \log \lambda - \sum \log x_i! \\ &= n\bar{x} \log \lambda - n\lambda + \text{constant},\end{aligned}$$

as required.

Question 3(c)

$$\frac{\partial \log f(\mathbf{x}|\lambda)}{\partial \lambda} = n\bar{x}\lambda^{-1} - n$$

$$\frac{\partial^2 \log f(\mathbf{x}|\lambda)}{\partial \lambda^2} = -n\bar{x}\lambda^{-2}$$

Therefore

$$\begin{aligned} I(\lambda) &= E_{\mathbf{X}|\lambda} \left[\frac{n\bar{x}}{\lambda^2} \right] \\ &= \frac{n}{\lambda^2} E_{\mathbf{X}|\lambda} \left[\frac{X_1 + \dots + X_n}{n} \right] \\ &= \frac{n}{\lambda^2} \frac{1}{n} (E_{\mathbf{X}|\lambda}[X_1] + \dots + E_{\mathbf{X}|\lambda}[X_n]) \\ &= \frac{1}{\lambda^2} (n\lambda) = \frac{n}{\lambda}. \end{aligned}$$

Question 3(c)

Therefore,

$$\pi(\lambda) \propto \sqrt{\frac{n}{\lambda}} \propto \frac{1}{\sqrt{\lambda}},$$

as required.

Question 3(d)

Let X : Number of road traffic accidents (per year).

$$f(\mathbf{x}|\lambda) = \frac{e^{-\lambda}\lambda^4}{4!} \times \frac{e^{-\lambda}\lambda^6}{6!} = \frac{e^{-2\lambda}\lambda^{10}}{17280}.$$

Thus

$$\pi(\lambda|\mathbf{x}) \propto \text{prior} \times \text{likelihood}$$

$$\propto \frac{1}{\sqrt{\lambda}} \times e^{-2\lambda}\lambda^{10}$$

$$\propto \lambda^{-\frac{1}{2}}\lambda^{10}e^{-2\lambda}$$

$$\propto \lambda^{9.5}e^{-2\lambda},$$

which is proportional to the pdf for a *Gamma*(10.5, 2), as required.

Question 3(e)

For large n , we know that

$$\lambda|\mathbf{x} \sim N\left(\hat{\lambda}, J(\hat{\lambda})^{-1}\right).$$

Now we find the MLE, $\hat{\lambda}$, by maximising the log-likelihood; thus

$$\frac{\partial}{\partial \lambda} \log f(\mathbf{x}|\lambda) = 0$$

gives

$$\frac{n\bar{x}}{\hat{\lambda}} = n, \quad \text{giving}$$

$$\hat{\lambda} = \bar{x}.$$

Question 3(e)

Now

$$J(\lambda) = -\frac{\partial^2}{\partial \lambda^2} \log f(\mathbf{x}|\lambda) = \frac{n\bar{x}}{\lambda^2}.$$

Thus

$$J(\hat{\lambda})^{-1} = \frac{\bar{x}^2}{n\bar{x}} = \frac{\bar{x}}{n},$$

as required.

Question 4(a)

Under certain regularity conditions

$$\mathbf{T}(\mathbf{X}) \text{ is sufficient for } \theta \iff f(\mathbf{x}|\theta) = h(\mathbf{x}) g(\mathbf{t}(\mathbf{x}), \theta)$$

for some functions h and g .

Question 4(b)

The likelihood is

$$\begin{aligned}f(\mathbf{x}|\theta) &= \prod_{i=1}^n \theta(1-\theta)^{x_i-1} \\&= \prod_{i=1}^n \theta(1-\theta)^{-1}(1-\theta)^{x_i} \\&= \left(\frac{\theta}{1-\theta}\right)^n \{(1-\theta)^{x_1} \times \dots \times (1-\theta)^{x_n}\} \\&= 1 \times \left(\frac{\theta}{1-\theta}\right)^n (1-\theta)^{\sum x_i} \\&= h(\mathbf{x}) \times g\left(\sum_{i=1}^n x_i, \theta\right),\end{aligned}$$

Question 4(b)

where

- $h(\mathbf{x}) = 1$ and

- $g(t, \theta) = \left(\frac{\theta}{1-\theta}\right)^n (1-\theta)^t,$

with $t = \sum_{i=1}^n x_i.$

Therefore, by the FT, $T = \sum_{i=1}^n X_i$ is sufficient for θ .

Question 5(a)(i)

The **bisection method** has been used. Here, the expert is asked to provide a plausible range of values for the parameter of interest θ , say $a \rightarrow b$.

Then, the expert is asked to bisect this range into two parts of equal probability, essentially eliciting the median for θ , say m .

Finally, the expert is asked to further bisect (a, m) and (m, b) into two parts of equal probability, giving elicited lower and upper quartiles.

A parametric distribution is then fitted to these judgements.

Question 5(a)(ii)

$$\theta \sim \text{Beta}(9, 2)$$

$$\pi(\theta) = \frac{\theta^8(1-\theta)}{B(9, 2)} \quad 0 \leq \theta \leq 1$$

$$E(\theta) = \frac{9}{11} = 0.818$$

$$SD(\theta) = \sqrt{\frac{9 \times 2}{(9+2)^2(9+2+1)}} = 0.111$$

Question 5(a)(iii)

The lower feedback percentile tells us that

$$\Pr(\theta < 0.502) = 0.01.$$

Thus, a suitable question might be:

“About 99 in a hundred Caribbean reefs have at least a fifty–fifty chance of observing a catastrophic decline in coral cover... does this seem plausible to you?”

Question 5(b)

We have a binomial likelihood:

$$f(\mathbf{x}|\theta) = {}^8C_6\theta^6(1-\theta)^2.$$

Combining this with the prior gives

$$\pi(\theta|\mathbf{x}) \propto \theta^8(1-\theta) \times \theta^6(1-\theta)^2$$

$$\propto \theta^{15-1}(1-\theta)^{4-1},$$

i.e. $\theta|x = 6 \sim \text{Beta}(15, 4)$.

Question 5(b)

Short report:

The posterior mean and standard deviation is:

$$E(\theta|x = 6) = \frac{15}{19} = 0.786$$

$$SD(\theta|x = 6) = \sqrt{\frac{60}{19^2 \times 20}} = 0.091.$$

Thus, in light of the data (which gives $\hat{\theta} = 0.75$), the beliefs of the biologist – concerning the proportion of coral reefs that are likely to experience a catastrophic reduction in live coral cover – have been shifted down slightly.

Including the data has also increased our estimation precision about θ , since the SD has reduced from 0.111 to 0.091.

Question 6(a)

We have substantial prior information for θ when the prior distribution dominates the posterior distribution, that is $\pi(\theta|\mathbf{x}) \sim \pi(\theta)$.

If we have vague prior knowledge, most of the information about θ will be passed through to the posterior distribution via the data, and so we have $\pi(\theta|\mathbf{x}) \sim f(\mathbf{x}|\theta)$.

Question 6(b)

A family $\mathfrak{F}$ of prior distributions for θ is said to be conjugate to $f(x|\theta)$ if for every prior distribution $\pi(\theta) \in \mathfrak{F}$, the posterior distribution $\pi(\theta|\mathbf{x})$ is also in $\mathfrak{F}$.

Question 6(c)(i)

If $X \sim \text{Ga}(4, \theta)$, then

$$\begin{aligned}f(\mathbf{x}|\theta) &= \prod_{i=1}^n \frac{\theta^4 x_i^3 e^{-\theta x_i}}{\Gamma(4)} \\&= \left(\frac{1}{\Gamma(4)}\right)^n \left\{(\theta^4 \dots \theta^4) \times (x_1^3 \dots x_n^3) \times (e^{-\theta x_1} \dots e^{-\theta x_n})\right\} \\&= \left(\frac{1}{6}\right)^n \left(\prod_{i=1}^n x_i^3\right) \theta^{4n} e^{-\theta \sum x_i} \\&= k \theta^{4n} e^{-n\bar{x}\theta}, \quad \text{as required.}\end{aligned}$$

Question 6(c)(ii)

$$\begin{aligned}\pi(\theta|\mathbf{x}) &\propto \pi(\theta) \times f(\mathbf{x}|\theta) \\ &\propto \theta^{g-1}e^{-h\theta} \times \theta^{4n}e^{-n\bar{x}\theta} \\ &\propto \theta^{(4n+g)-1}e^{-(h+n\bar{x})\theta},\end{aligned}$$

so $\theta|\mathbf{x} \sim \text{Ga}(G = 4n + g, H = h + n\bar{x})$.

Question 6(c)(iii)

The gamma distribution *is* a conjugate prior here – the posterior is in the same family as the prior (both gamma distributions).

Question 6(c)(iv)

We represent vague prior knowledge by using a conjugate prior with very large variance.

If we have a $Ga(g, h)$ prior for θ , then

$$E(\theta) = m = \frac{g}{h} \quad \text{and} \quad \text{Var}(\theta) = v = \frac{g}{h^2}.$$

Rearranging, we get

$$g = \frac{m^2}{v} \quad \text{and} \quad h = \frac{m}{v},$$

and so clearly, as $v \rightarrow \infty$, both $g, h \rightarrow 0$.

Question 7(a)

Applying Bayes Theorem, the posterior density function is

$$\pi(\mu|\mathbf{x}) \propto \pi(\mu)f(\mathbf{x}|\mu)$$

$$\begin{aligned} &\propto \left(\frac{d}{2\pi}\right)^{1/2} \exp\left\{-\frac{d}{2}(\mu - b)^2\right\} \\ &\quad \times \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left\{-\frac{n\tau}{2}\left[s^2 + (\bar{x} - \mu)^2\right]\right\} \end{aligned}$$

$$= k_1 \exp\left\{-\frac{d}{2}(\mu - b)^2 - \frac{n\tau}{2}\left[s^2 + (\bar{x} - \mu)^2\right]\right\}$$

$$= k_1 \exp\left\{-\frac{1}{2}\left[d(\mu - b)^2 + n\tau(\bar{x} - \mu)^2\right]\right\}$$

where k_1 is a constant that does not depend on μ .

Question 7(a)

We can also simplify the exponent. We have

$$\begin{aligned}d(\mu - b)^2 &+ n\tau(\bar{x} - \mu)^2 \\&= d(\mu^2 - 2b\mu + b^2) + n\tau(\bar{x}^2 - 2\bar{x}\mu + \mu^2) \\&= d\mu^2 - 2db\mu + db^2 + n\tau\bar{x}^2 - 2n\tau\bar{x}\mu + n\tau\mu^2 \\&= (d + n\tau)\mu^2 - 2\mu(db + n\tau\bar{x}) + c \\&= (d + n\tau) \left(\mu^2 - \frac{2(db + n\tau\bar{x})}{d + n\tau} \mu + \frac{c}{d + n\tau} \right) \\&= (d + n\tau) \left\{ \mu - \left(\frac{db + n\tau\bar{x}}{d + n\tau} \right) \right\}^2 + c\end{aligned}$$

where c does not depend on μ .

Question 7(a)

Let

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

Then

$$\begin{aligned}\pi(\mu|\mathbf{x}) &= k_1 \exp \left\{ -\frac{D}{2}(\mu - B)^2 - \frac{c}{2} \right\} \\ &= k \exp \left\{ -\frac{D}{2}(\mu - B)^2 \right\},\end{aligned}$$

where k is a constant that does not depend on μ .

Thus $\mu|\mathbf{x} \sim N(B, 1/D)$.

Question 7(b)(i)

A 95% Bayesian confidence interval is any interval C_B with

$$\Pr(\mu \in C_B | \mathbf{x}) = 0.95.$$

Question 7(b)(ii)

The HDI for μ contains the most likely values of μ and, as the posterior distribution is symmetric about B with variance $1/D$, the HDI is an equi-tailed interval:

$$C_H = \left(B - 1.96\sqrt{\frac{1}{D}}, B + 1.96\sqrt{\frac{1}{D}} \right).$$

Question 7(b)(iii)

The CLT tells us that

$$\bar{X} \sim N\left(\mu, \frac{1/\tau}{n}\right), \quad \text{approximately,}$$

and so the frequentist equi-tailed interval is

$$C_F = \left(\bar{x} - 1.96\sqrt{\frac{1}{n\tau}}, \bar{x} + 1.96\sqrt{\frac{1}{n\tau}} \right).$$

Question 7(b)(iv)

- The probability that C_B and C_H contain μ is 0.95
- The probability that C_F contain μ is either 0 or 1, since μ does not have a probability distribution; rather, it is a fixed (but unknown) constant.
- Unlike C_H , C_B might not contain the most likely values of μ .

Question 8

Using Bayes' Theorem, the posterior density for θ given $\mathbf{x}$ and y is

$$\begin{aligned}\pi(\theta|\mathbf{x}, y) &= \frac{\pi(\theta)f(\mathbf{x}, y|\theta)}{f(\mathbf{x}, y)} \\&= \frac{\pi(\theta)f(\mathbf{x}|\theta)f(y|\theta)}{f(\mathbf{x})f(y|\mathbf{x})} \quad \text{since } \mathbf{X} \text{ and } Y \text{ are independent} \\&= \frac{\pi(\theta|\mathbf{x})f(y|\theta)}{f(y|\mathbf{x})}.\end{aligned}$$

Question 8

Rearranging, we obtain

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