1. Problem 10.1 A forecaster must announce probabilities $q = (q_1, \ldots, q_J)$ for the events $\theta_1, \ldots, \theta_J$. These events form a partition: that is, one and only one of them will occur. The forecaster will be scored based on the scoring rule

$$s(\theta_j, q) = \sum_{i=1}^{J} |q_i - 1_{i=1}|.$$

Let $\pi = (\pi_1, \ldots, \pi_J)$ the forecaster’s own probability for the events on events $\theta_1, \ldots, \theta_J$. Show that this scoring rule is not proper. Because you are looking for a counterexample, it is okay to consider a simplified version of the problem, for example by picking a small $J$.

Proof. We need to find a vector $q \neq \pi$ such that

$$\sum_{j=1}^{J} s(\theta_j, q)\pi_j < \sum_{j=1}^{J} s(\theta_j, \pi)\pi_j.$$

Suppose $J = 2$. Let $q_1 = q$, $q_2 = 1 - q$. Also let $q_1 = q$, $q_2 = 1 - q$, $\pi_1 = \pi$, and $\pi_2 = 1 - \pi$. Then

$$s(\theta_1, q) = |q_1 - 1_{1=1}| + |q_2 - 1_{2=1}|$$

$$= |q - 1| + (1 - q)$$

$$= 2(1 - q),$$

and

$$s(\theta_2, q) = |q_1 - 1_{1=2}| + |q_2 - 1_{2=2}|$$

$$= q + |1 - q - 1|$$

$$= 2q.$$

Therefore,

$$\sum_{j=1}^{J} s(\theta_j, q)\pi_j = s(\theta_1, q)\pi + s(\theta_2, q)(1 - \pi)$$

$$= 2(1 - q)\pi + 2q(1 - \pi)$$

$$= 2\pi + (2 - 4\pi)q.$$
If \( \pi > 0.5 \), \( \pi = 0.75 \) for example, \( \sum_{j=1}^{J} s(\theta_j, q)\pi_j \) is minimized at \( q = 1 \) and \( \min_q \sum_{j=1}^{J} s(\theta_j, q)\pi_j = 0.5 \). If \( q = \pi = 0.75 \), \( \sum_{j=1}^{J} s(\theta_j, \pi)\pi_j = 0.75 \). We show that \( \sum_{j=1}^{J} s(\theta_j, q)\pi_j < \sum_{j=1}^{J} s(\theta_j, \pi)\pi_j \). Hence \( s(\theta_j, q) = \sum_{i=1}^{J} |q_i - 1|_{i=j} \) is not proper.

2. Problem 10.6 Consider a sequence of independent binary events with probability of success 0.4. Evaluate the two terms in equation (10.8) for the following four forecasters:

Charles: always says 0.4.
Mary: randomly chooses between 0.3 and 0.5.
Qing: says either 0.2 or 0.3; when he says 0.2 it never rains, when he says 0.3 it always rains.
Ana: follows this table:

<table>
<thead>
<tr>
<th>( \pi )</th>
<th>Rain</th>
<th>No Rain</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.3</td>
<td>0.15</td>
<td>0.35</td>
</tr>
<tr>
<td>0.5</td>
<td>0.25</td>
<td>0.25</td>
</tr>
</tbody>
</table>

Comment on the calibration and refinement of these forecasters.

Solution:
Suppose we have a sequence \( x_1, \ldots, x_n \) of independent binary events with probability of success 0.4, i.e., \( \bar{x} = 0.4 \). We can view this as relative frequency of rain among \( n \) days, for example.

The equation (10.8) is

\[
BS = \sum_{\pi \in \Pi} \nu(\pi)[\pi - \bar{x}(\pi)]^2 + \sum_{\pi \in \Pi} \nu(\pi)[\bar{x}(\pi)(1 - \bar{x}(\pi))].
\]

The first term on the right hand side is a measure of calibration of the forecaster, while the second term measures the refinement.

- Charles always says \( \pi = 0.4 \).
  
The Brier score is
  
  \[
  BS = (0.4 - 0.4)^2 + (0.4)(1 - 0.4) = 0 + 0.24 = 0.24.
  \]
  
  Since Charles’s forecast always matches the relative frequency of rain, 0.4, it is perfectly calibrated and so the first term of the BS score is zero. It’s sharpness is not quite good since \( \bar{x} \) is not close to zero or one, but close to the least refinement case that the relative frequency is 0.5.

- Mary randomly chooses between 0.3 and 0.5.
  
  Since Mary chooses her forecast either 0.3 or 0.5 at random, the \( \nu(\pi = 0.3) \) and \( \mu(\pi = 0.5) \) should be 0.5 or very close to 0.5. Assume both are 0.5. Also, both \( \bar{x}(\pi = 0.3) \) and \( \bar{x}(\pi = 0.5) \) should be equal or very close to 0.4. Assume \( \bar{x}(\pi = 0.3) = \bar{x}(\pi = 0.5) = 0.4 \).
  
  Then the BS score is
  
  \[
  BS = (0.5)[(0.3 - 0.4)^2 + (0.4)(1 - 0.4)] + (0.5)[(0.5 - 0.4)^2 + (0.4)(1 - 0.4)] = 0.01 + 0.24 = 0.25.
  \]
  
  Since Mary can’t predict 0.4 precisely, her forecast is not as well calibrated as Charles. Its refinement is not improved because of choosing 0.3 or 0.5 randomly without any strategy.
When Qing says 0.2 it never rains, when he says 0.3 it always rains.

From Qing’s forecast, we know that \( \nu(0.2) = 0.6, \nu(0.3) = 0.4, \bar{x}(0.2) = 0 \) and \( \bar{x}(0.3) = 1 \). Hence the Brier score is

\[
BS = (0.6)(0.2 - 0)^2 + (0.4)(0.3 - 1)^2 + (0.6)(0(1 - 0)) + (0.4)(1(1 - 1)) \\
= 0.22 + 0 = 0.22.
\]

Qing’s forecast is not well calibrated because both 0.2 and 0.3 are away from their corresponding relative frequency, 0 and 1. However, his forecast perfectly separates Rain and No Rain groups, and hence has the best refinement (perfect sharpness).

According to Ana’s forecast, \( \nu(0.3) = \nu(0.5) = 0.5, \) and \( \bar{x}(0.3) = 0.3 \) and \( \bar{x}(0.5) = 0.5 \). The Brier score is

\[
BS = (0.5)(0.3 - 0.3)^2 + (0.5)(0.5 - 0.5)^2 + (0.5)((0.3)(1 - 0.3)) + (0.5)((0.5)(1 - 0.5)) \\
= 0 + 0.23 = 0.23.
\]

Her forecast is well calibrated but both \( \pi = 0.3 \) and \( \pi = 0.5 \) cases contain a few rain and no-rain events, and so it’s not good for refinement.

3. **Problem 13.3** Using simulation, approximate the distribution of \( V_x(E) \) in the example in Section 13.1.4. Suppose each observation costs \$230. Compute the marginal cost-effectiveness ratio for performing the experiment in the example (versus the status quo of no experimentation). Use the negative of the loss as the measure of effectiveness.

**Solution:**

Without loss of generality and for simplicity, assume the following hyperparameters: \( \mu_0 = 0, \tau_0^2 = 4, \sigma^2 = 1 \). To approximate the distribution of \( V_x(E) \), I generate observed values of size 1000 from the prior predictive distribution \( x \sim N(\mu_0, \sigma^2 + \tau_0^2) \), and then for each \( x \), compute its corresponding posterior mean \( E(\theta|x) \), and get the 1000 different \( V_x(E) \)'s via the fact \( V_x(E) = (E(\theta|x) - E(\theta))^2 \). Figure 1 shows the approximated \( V_x(E) \) and \( \frac{\sigma^2 + \tau_0^2}{\tau_0^2} V_x(E) \).

Note that the distribution of \( \frac{\sigma^2 + \tau_0^2}{\tau_0^2} V_x(E) \) is \( \chi^2 \) (red curve).

Now suppose the true value of \( \theta \) is 3. Suppose squared loss is used and the negative of the loss is the measure of effectiveness, i.e., \(- (\theta - a)^2 \). Let marginal cost-effectiveness ratio be defined as \(- 230 / (\theta - a)^2 \) if we do the experiment, 0 if we don’t do experiment. Suppose we do the experiment and one observation \( x \) is drawn and has value 2.439524. The posterior mean is

\[
\mu_x = \frac{\sigma^2}{\sigma^2 + \tau_0^2} \mu_0 + \frac{\tau_0^2}{\sigma^2 + \tau_0^2} x = \frac{1}{1 + 4}(0) + \frac{4}{1 + 4}(2.439524) = 1.951619
\]

Then the cost-effectiveness ratio is \(- 230 / (3 - 1.951619)^2 = - 209.2616 \). We are hoping the ratio to be small because it means that the drawn \( x \) has value closer to the true value of \( \theta \). But it is kind of useless because we don’t know the value of \( \theta \). Probably we need another measure that does not involve unknown values.

4. **Problem 13.5** Consider an experiment consisting of a single Bernoulli observation, from a population with success probability \( \theta \). Consider a simple versus simple hypothesis testing situation in which \( A = \{a_0, a_1\}, \Theta = \{0.50, 0.75\} \), and with utilities as shown in Table 13.6. Compute \( V(E) = E_x[V(\pi_x)] - V(\pi) \).

**Solution:**

The table 13.6 is
Figure 1: Approximated $V_\delta(\mathcal{E})$ and $\frac{\sigma^2 + \tau_0^2}{\tau_0} V_\delta(\mathcal{E})$
The hypothesis is $H_0 : \theta = 0.5$ and $H_1 : \theta = 0.75$. Set $\pi(H_0) = \pi(H_1) = 1/2$. First find the Bayes action with respect to the prior $\pi$.

$$U_\pi(a_0) = u(a_0(H_0))\pi(H_0) + u(a_0(H_1))\pi(H_1)$$

$$= 0(1/2) - 1(1/2) = -1/2$$

$$U_\pi(a_1) = u(a_1(H_0))\pi(H_0) + u(a_1(H_1))\pi(H_1)$$

$$= -3(1/2) + 0(1/2) = -3/2$$

Hence $a^* = a_0$ and $V(\pi) = \sup_{a \in \mathcal{A}} U_\pi(a) = -1/2$.

Let the prior for $\theta$ be $\theta \sim \text{Beta}(\alpha, \beta)$. Then for a single observation $x$, the posterior distribution is

$$\pi(\theta|x) \propto \theta^{x+\alpha-1}(1-\theta)^{1-x+\beta-1} \sim \text{Beta}(x+\alpha, 1-x+\beta)$$

Then the posterior density of $H_0$ and $H_1$ are

$$\pi(H_0|x) = \pi(\theta = 0.5|x) = \frac{\Gamma(\alpha + \beta + 1)}{\Gamma(x + \alpha)\Gamma(1-x + \beta)}(0.5)^{x+\alpha-1}(0.5)^{1-x+\beta-1}$$

$$= \frac{\Gamma(\alpha + \beta + 1)}{\Gamma(x + \alpha)\Gamma(1-x + \beta)}(0.5)^{\alpha+\beta-1}$$

$$\pi(H_1|x) = \pi(\theta = 0.75|x) = \frac{\Gamma(\alpha + \beta + 1)}{\Gamma(x + \alpha)\Gamma(1-x + \beta)}(0.75)^{x+\alpha-1}(0.25)^{1-x+\beta-1}$$

Hence

$$U_{x_\pi}(a_0) = u(a_0(H_0))\pi_x(H_0) + u(a_0(H_1))\pi_x(H_1)$$

$$= 0\pi(H_0|x) - 1\pi(H_1|x) = -\frac{\Gamma(\alpha + \beta + 1)}{\Gamma(x + \alpha)\Gamma(1-x + \beta)}(0.75)^{x+\alpha-1}(0.25)^{1-x+\beta-1}$$

$$U_{x_\pi}(a_1) = u(a_1(H_0))\pi_x(H_0) + u(a_1(H_1))\pi_x(H_1)$$

$$= -3\pi(H_0|x) + 0\pi(H_1|x) = (-3\frac{\Gamma(\alpha + \beta + 1)}{\Gamma(x + \alpha)\Gamma(1-x + \beta)}(0.5)^{\alpha+\beta-1}$$

If $x = 0$,

$$U_{x_\pi}(a_0) = -\frac{\Gamma(\alpha + \beta + 1)}{\Gamma(\alpha)\Gamma(1+ \beta)}(0.75)^{\alpha-1}(0.25)^\beta$$

$$U_{x_\pi}(a_1) = (-3\frac{\Gamma(\alpha + \beta + 1)}{\Gamma(\alpha)\Gamma(1+ \beta)}(0.5)^{\alpha+\beta-1}.$$

Assume $\beta = 1$. Then $U_{x_\pi}(a_0) > U_{x_\pi}(a_1)$ if and only if $\alpha < 5.4190$. We have to put lots of mass on $\theta$ close to 1, $\alpha > 5.4190$ so that the optimal choice becomes $a_1$. 

<table>
<thead>
<tr>
<th>Actions</th>
<th>States of Nature</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_0 = \text{accept } H_0$</td>
<td>$H_0$</td>
</tr>
<tr>
<td>$a_1 = \text{accept } H_1$</td>
<td>$-3$</td>
</tr>
</tbody>
</table>
If $x = 1$,
\[
U_{\pi_x}(a_0) = -\frac{\Gamma(\alpha + \beta + 1)}{\Gamma(\alpha + 1)\Gamma(\beta)} (0.75)\alpha (0.25)^{\beta - 1}
\]
\[
U_{\pi_x}(a_1) = (3) \frac{\Gamma(\alpha + \beta + 1)}{\Gamma(\alpha + 1)\Gamma(\beta)} (0.5)^{\alpha + \beta - 1}
\]

In this case, again with $\beta = 1$, $U_{\pi_x}(a_0) > U_{\pi_x}(a_1)$ if and only if $\alpha < 2.7095$. Since $x = 1$ is an evidence for larger value of $\theta$, $\alpha$ need not be so large to favor $a_1$.

In sum, with $\beta = 1$, we have three cases.

**CASE 1:** $\alpha < 2.7095$

$a_0$ is optimal for both $x = 0$ and $x = 1$, and

\[
V(\pi_x) = \max(U_{\pi_x}(a_0), U_{\pi_x}(a_1)) = \begin{cases} 
U_{\pi_x}(a_0) = -\frac{\Gamma(\alpha + 2)}{\Gamma(\alpha)\Gamma(2)} (0.75)\alpha - 1(0.25) & \text{if } x = 0 \\
U_{\pi_x}(a_0) = -\frac{\Gamma(\alpha + 2)}{\Gamma(\alpha + 1)\Gamma(1)} (0.75)^{\alpha} & \text{if } x = 1
\end{cases}
\]

Therefore, the observed information is

\[
V_x(\mathcal{E}) = V(\pi_x) - U_{\pi_x}(a_0) = 0
\]

for $x = 0, 1$.

Hence $V(\mathcal{E}) = E_x[V(\pi_x)] - V(\pi) = 0$

**CASE 2:** $2.7095 < \alpha < 5.4190$

In this case,

\[
V(\pi_x) = \begin{cases} 
U_{\pi_x}(a_0) = -\frac{\Gamma(\alpha + 2)}{\Gamma(\alpha)\Gamma(2)} (0.75)\alpha - 1(0.25) & \text{if } x = 0 \\
U_{\pi_x}(a_1) = (3) \frac{\Gamma(\alpha + 2)}{\Gamma(\alpha + 1)\Gamma(1)} (0.5)^{\alpha} & \text{if } x = 1
\end{cases}
\]

Therefore, the observed information is

\[
V_x(\mathcal{E}) = V(\pi_x) - U_{\pi_x}(a_0) = \begin{cases} 
0 & \text{if } x = 0 \\
(-3) \frac{\Gamma(\alpha + 2)}{\Gamma(\alpha + 1)\Gamma(1)} (0.5)^{\alpha} + \frac{\Gamma(\alpha + 2)}{\Gamma(\alpha + 1)\Gamma(1)} (0.75)^{\alpha - 1} & \text{if } x = 1
\end{cases}
\]

The marginal of $x$ is

\[
m(x) = \int f(x|\theta)\pi(\theta)d\theta
\]

\[
= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \int \theta^{x+\alpha-1}(1 - \theta)^{1-x+\beta-1}d\theta
\]

\[
= \frac{\Gamma(\alpha + \beta) \Gamma(x + \alpha)\Gamma(1 - x + \beta)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\alpha + \beta + 1)}
\]

\[
= \frac{\Gamma(x + \alpha)\Gamma(1 - x + \beta)}{\Gamma(\alpha)\Gamma(\beta)(\alpha + \beta)}
\]

If $\beta = 1$, then

\[
m(x) = \frac{\Gamma(x + \alpha)\Gamma(2 - x)}{\Gamma(\alpha)(\alpha + 1)}
\]
Hence

\[ V(\mathcal{E}) = E_x[V(\pi_x)] - V(\pi) = \left[ -3 \frac{\Gamma(\alpha + 2)}{\Gamma(\alpha + 1)\Gamma(1)} (0.5)^\alpha + \frac{\Gamma(\alpha + 2)}{\Gamma(\alpha + 1)\Gamma(1)} (0.75)^{\alpha - 1} \right] \frac{\Gamma(1 + \alpha)\Gamma(1)}{\Gamma(\alpha)\Gamma(\alpha + 1)} \]

\[ = (-3)\alpha(0.5)^\alpha + \alpha(0.75)^{\alpha - 1} \]

**CASE 3:** \( \alpha > 5.4190 \)

In this case,

\[ V(\pi_x) = \begin{cases} 
U_x(a_1) = (-3) \frac{\Gamma(\alpha + 2)}{\Gamma(\alpha)\Gamma(2)} (0.5)^\alpha & \text{if } x = 0 \\
U_x(a_1) = (-3) \frac{\Gamma(\alpha + 2)}{\Gamma(\alpha + 1)\Gamma(1)} (0.5)^\alpha & \text{if } x = 1 
\end{cases} \]

\[ V_x(\mathcal{E}) = V(\pi_x) - U_x(a_0) = \begin{cases} 
(-3) \frac{\Gamma(\alpha + 2)}{\Gamma(\alpha)\Gamma(2)} (0.5)^\alpha + \frac{\Gamma(\alpha + 2)}{\Gamma(\alpha + 1)\Gamma(1)} (0.75)^{\alpha - 1}(0.25) & \text{if } x = 0 \\
(-3) \frac{\Gamma(\alpha + 2)}{\Gamma(\alpha + 1)\Gamma(1)} (0.5)^\alpha + \frac{\Gamma(\alpha + 2)}{\Gamma(\alpha + 1)\Gamma(1)} (0.75)^{\alpha - 1} & \text{if } x = 1 
\end{cases} \]

Hence

\[ V(\mathcal{E}) = E_x[V(\pi_x)] - V(\pi) = \left[ -3 \frac{\Gamma(\alpha + 2)}{\Gamma(\alpha)\Gamma(2)} (0.5)^\alpha + \frac{\Gamma(\alpha + 2)}{\Gamma(\alpha)\Gamma(2)} (0.75)^{\alpha - 1}(0.25) \right] \frac{\Gamma(\alpha)\Gamma(2)}{\Gamma(\alpha)\Gamma(\alpha + 1)} \]

\[ + \left[ -3 \frac{\Gamma(\alpha + 2)}{\Gamma(\alpha + 1)\Gamma(1)} (0.5)^\alpha + \frac{\Gamma(\alpha + 2)}{\Gamma(\alpha + 1)\Gamma(1)} (0.75)^{\alpha - 1} \right] \frac{\Gamma(1 + \alpha)\Gamma(1)}{\Gamma(\alpha)\Gamma(\alpha + 1)} \]

\[ = (-3)\alpha(0.5)^\alpha + \alpha(0.75)^{\alpha - 1}(0.25) + (-3)\alpha(0.5)^\alpha + \alpha(0.75)^{\alpha - 1} \]

\[ = (-6)\alpha(0.5)^\alpha + \alpha(0.5)^{\alpha - 1}(1.25) \]

5. **Problem 13.11** Consider an experiment \( \mathcal{E} \) that consists of observing \( n \) conditionally independent random variables \( x_1, \ldots, x_n \), with \( x_i \sim N(\theta, \sigma^2) \), with \( \sigma \) known. Suppose also that a priori \( \theta \sim N(\mu, \tau^2) \). Show that

\[ I(\mathcal{E}) = \frac{1}{2} \log \left( 1 + n \frac{\tau^2}{\sigma^2} \right) \]

**Proof.** From the problem, we know the marginal of \( \bar{x} \) is also normally distributed. With the conditional expectation and conditional variance identities, we have

\[ \bar{x} | \theta \sim N \left( \theta, \frac{\sigma^2}{n} \right) \]

\[ \bar{x} \sim N \left( \mu, \frac{\sigma^2}{n} + \tau^2 \right) \]

Hence

\[ f(\bar{x} | \theta) = \left( \frac{n}{2\pi\sigma^2} \right)^{1/2} \exp \left\{ -\frac{n}{2\sigma^2}(\bar{x}^2 - 2\bar{x}\theta + \theta^2) \right\} \]

\[ m(\bar{x}) = \left( \frac{n}{2\pi(\sigma^2 + n\tau^2)} \right)^{1/2} \exp \left\{ -\frac{n}{2(\sigma^2 + n\tau^2)}(\bar{x}^2 - 2\bar{x}\mu + \mu^2) \right\} . \]
Taking log, we have
\[
\log f(\bar{x}|\theta) = \frac{1}{2} \left[ \log \left( \frac{n}{2\pi} \right) - \log(\sigma^2) \right] - \frac{n}{2\sigma^2} (\bar{x}^2 - 2\bar{x}\theta + \theta^2)
\]
\[
\log m(\bar{x}) = \frac{1}{2} \left[ \log \left( \frac{n}{2\pi} \right) - \log(\sigma^2 + n\tau^2) \right] - \frac{n}{2(\sigma^2 + n\tau^2)} (\bar{x}^2 - 2\bar{x}\mu + \mu^2).
\]

Therefore,
\[
\log \left( \frac{f(\bar{x}|\theta)}{m(\bar{x})} \right) = \frac{1}{2} \left[ \log(\sigma^2 + n\tau^2) - \log(\sigma^2) \right] - \frac{n}{2\sigma^2} (\bar{x}^2 - 2\bar{x}\theta + \theta^2) + \frac{n}{2(\sigma^2 + n\tau^2)} (\bar{x}^2 - 2\bar{x}\mu + \mu^2)
\]

Since \( E_{x|\theta}[\bar{x}] = \theta \) and \( E_{x|\theta}[\bar{x}^2] = \frac{\sigma^2}{n} + \theta^2 \),
\[
E_{x|\theta}[A] = -\frac{n}{2\sigma^2} (E_{x|\theta}[\bar{x}^2] - 2E_{x|\theta}[\bar{x}]/\theta + \theta^2) + \frac{n}{2(\sigma^2 + n\tau^2)} (E_{x|\theta}[\bar{x}^2] - 2E_{x|\theta}[\bar{x}]\mu + \mu^2)
\]
\[
= -\frac{n}{2\sigma^2} \left( \frac{\sigma^2}{n} + \theta^2 - 2\theta^2 + \theta^2 \right) + \frac{n}{2(\sigma^2 + n\tau^2)} \left( \frac{\sigma^2}{n} + \theta^2 - 2\theta^2 + \theta^2 \right)
\]
\[
= -\frac{n}{2}\frac{\sigma^2}{\sigma^2 + n\tau^2} + \frac{n}{2}\frac{\sigma^2}{\sigma^2 + n\tau^2}
\]
\[
= 0
\]

Then, with \( E_{\theta}[\theta] = \mu \) and \( E_{\theta}[\theta^2] = \tau^2 + \mu^2 \), we have
\[
E_{\theta} \left[ E_{x|\theta}[A] \right] = -\frac{1}{2} + \frac{n}{2(\sigma^2 + n\tau^2)} \left( \frac{\sigma^2}{n} + E_{\theta}[\theta^2] - 2E_{\theta}[\theta]\mu + \mu^2 \right)
\]
\[
= -\frac{1}{2} + \frac{n}{2(\sigma^2 + n\tau^2)} \left( \frac{\sigma^2}{n} + \tau^2 + \mu^2 - 2\mu\mu + \mu^2 \right)
\]
\[
= -\frac{1}{2} + \frac{1}{2}
\]
\[
= 0
\]

Thus, since \( \frac{1}{2} [\log(\sigma^2 + n\tau^2) - \log(\sigma^2)] \) is constant of \( x|\theta \) and \( \theta \), the Lindley information is
\[
\mathcal{I}(\mathcal{E}) \equiv E_{\theta} \left[ E_{x|\theta} \left[ \log \left( \frac{f(\bar{x}|\theta)}{m(\bar{x})} \right) \right] \right]
\]
\[
= \frac{1}{2} \left[ \log(\sigma^2 + n\tau^2) - \log(\sigma^2) \right] + E_{\theta} \left[ E_{x|\theta}[A] \right]
\]
\[
= \frac{1}{2} \log \left( 1 + \frac{n\tau^2}{\sigma^2} \right) + 0
\]
\[
= \frac{1}{2} \log \left( 1 + \frac{n\tau^2}{\sigma^2} \right).
\]

This completes the proof.