1 Problem 1.4

In order to find an explicit form of the integral
\[
\int_{\omega}^{\infty} \alpha \beta x^{\alpha - 1} e^{-\beta x} \, dx,
\]
we use the change of variable \( y = x^\alpha \). We have \( dy = \alpha x^{\alpha - 1} \, dx \) and the integral becomes
\[
\int_{\omega}^{\infty} \alpha \beta x^{\alpha - 1} e^{-\beta x} \, dx = \int_{\omega}^{\infty} \beta e^{-\beta y} \, dy = e^{-\beta \omega}.
\]

2 Problem 1.7

The density \( f \) of the vector \( Y \) is
\[
f(y_n, \mu, \sigma) = \left( \frac{1}{\sigma \sqrt{2\pi}} \right)^n \exp \left( -\frac{1}{2} \sum_{i=1}^{n} \left( \frac{y_i - \mu}{\sigma} \right)^2 \right), \quad \forall y_n \in \mathbb{R}^n, \forall (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}^+.
\]
This function is strictly positive and the first and second order partial derivatives with respect to \( \mu \) and \( \sigma \) exist and are positive. The same hypotheses are satisfied for the log-likelihood function
\[
\log(L(\mu, \sigma, y_n)) = -n \log \sqrt{2\pi} - n \log \sigma - \frac{1}{2} \sum_{i=1}^{n} \left( \frac{y_i - \mu}{\sigma} \right)^2
\]
thus we can find the ML estimator of \( \mu \) and \( \sigma^2 \). The gradient of the log-likelihood is
\[
\nabla \log(L) = \left\{ \begin{array}{ll}
\frac{\partial \log(L(\mu, \sigma, y_n))}{\partial \mu} & = \frac{1}{\sigma} \sum_{i=1}^{n} (y_i - \mu) \\
\frac{\partial \log(L(\mu, \sigma, y_n))}{\partial \sigma} & = -\frac{n}{\sigma} + \sum_{i=1}^{n} (y_i - \mu)^2 / \sigma^3
\end{array} \right.
\]
if we equate the gradient to the null vector, \( \nabla \log(L) = 0 \) and solve the resulting system in \( \mu \) and \( \sigma \), we find
\[
\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} y_i = \bar{y},
\]
\[
\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (y_i - \bar{y})^2 = s^2.
\]

3 Problem 1.13

The function to be maximized in the most general case is
\[
\max_{c, \gamma, \alpha} c \prod_{i=1}^{17} \frac{c (x_i - \gamma)^{c-1} e^{-(x_i - \gamma)^c}}{\alpha^c (216 - \gamma)^c} \left( 1 - e^{-\frac{(216 - \gamma)^c}{\alpha^c}} \right) \left( 1 - e^{-\frac{(244 - \gamma)^c}{\alpha^c}} \right)
\]
where the first 17 \( x_i \)'s are the uncensored observations. Estimates are about .0078 for \( c \) and 3.5363 for \( \alpha \) when \( \gamma \) is equal to 100.
4 Problem 1.22

(a). Since $L(\delta, h(\theta)) \geq 0$ by using Fubini’s theorem, we get

\[
r(\pi, \delta) = \int_{\Theta} \int_{X} L(\delta, h(\theta)) f(x|\theta) \pi(\theta) d\theta dx
= \int_{X} \int_{\Theta} L(\delta, h(\theta)) f(x|\theta) \pi(\theta) d\theta dx
= \int_{X} \int_{\Theta} L(\delta, h(\theta)) m(x) \pi(\theta|x) d\theta dx
= \int_{X} \varphi(\pi, \delta|x) m(x) dx,
\]

where $m$ is the marginal distribution of $X$ and $\varphi(\pi, \delta|x)$ is the posterior average cost.

The estimator that minimizes the integrated risk $r$ is therefore, for each $x$, the one that minimizes the posterior average cost and it is given by

$$\delta^*(x) = \arg \min_{\delta} \varphi(\pi, \delta|x).$$

(b). The average posterior loss is given by:

\[
\varphi(\pi, \delta|x) = \mathbb{E}^\pi [L(\delta, \theta)|x]
= \mathbb{E}^\pi [|h(\theta) - \delta|^2|x]
= \mathbb{E}^\pi [|h(\theta)|^2|x] + \delta^2 - 2 < \delta, \mathbb{E}^\pi [h(\theta)|x] >
\]

A simple derivation shows that the minimum is attained for

$$\delta^*(x) = \mathbb{E}^\pi [h(\theta)|x] .$$

(c). Take $m$ to be the posterior median and consider the auxiliary function of $\theta$, $s(\theta)$, defined as

$$s(\theta) = \begin{cases} 
-1 & \text{if } h(\theta) < m \\
+1 & \text{if } h(\theta) > m 
\end{cases}$$

Then $s$ satisfies the propriety

\[
\mathbb{E}^\pi [s(\theta)|x] = -\int_{-\infty}^{m} \pi(\theta|x) d\theta + \int_{m}^{\infty} \pi(\theta|x) d\theta
= -\mathbb{P}(h(\theta) < m|x) + \mathbb{P}(h(\theta) > m|x) = 0
\]

For $\delta < m$, we have $L(\delta, \theta) - L(m, \theta) = |h(\theta) - \delta| - |h(\theta) - m|$ from which it follows that

\[
L(\delta, \theta) - L(m, \theta) = \begin{cases} 
\delta - m = (m - \delta)s(\theta) & \text{if } \delta > h(\theta) \\
m - \delta = m - \delta & \text{if } m < \delta \\
2h(\theta) - (\delta + m) > (m - \delta)s(\theta) & \text{if } \delta < h(\theta) < m
\end{cases}
\]
It turns out that \( L(\delta, \theta) - L(m, \theta) > (m - \delta)s(\theta) \) which implies that

\[
\mathbb{E}^\pi[ L(\delta, \theta) - L(m, \theta)|x] > (m - \delta)\mathbb{E}^\pi[s(\theta)|x] = 0.
\]

This still holds, using similar argument when \( \delta > m \), so the minimum of \( \mathbb{E}^\pi[ L(\delta, \theta)|x] \) is reached at \( \delta = m \).

5 Problem 1.23

(a). When \( X|\sigma \sim N(0, \sigma^2) \), \( \frac{1}{\sigma^2} \sim Ga(1, 2) \), the posterior distribution is

\[
\pi\left(\sigma^{-2}|X\right) \propto f(x|\sigma)\pi(\sigma^{-2}) \\
\propto \frac{1}{\sigma}e^{-\frac{(x^2/2+2)}{\sigma^2}} \\
= \left(\sigma^2\right)^{-1}e^{-\frac{(x^2/2+2)}{\sigma^2}},
\]

which means that \( 1/\sigma^2 \sim Ga\left(\frac{3}{2}, 2 + \frac{x^2}{2}\right) \). The marginal distribution is

\[
m(x) = \int f(x|\sigma)\pi(\sigma^{-2})d(\sigma^{-2}) \propto \left(\frac{x^2}{2} + 2\right)^{-\frac{3}{2}},
\]

that is, \( X \sim T(2, 0, 2) \).

(b). When \( X|\lambda \sim P(\lambda) \), \( \lambda \sim Ga(2, 1) \), the posterior distribution is

\[
\pi(\lambda) \propto f(x|\lambda)\pi(\lambda) \propto \lambda x^1 e^{-2\lambda}
\]

which means that \( \lambda \sim Ga(x + 2, 2) \). The marginal distribution is

\[
m(x) = \int f(x|\lambda)\pi(\lambda)d\lambda \propto \frac{\Gamma(x + 2)}{\sqrt{\pi}2^{x+2}x!} = \frac{(x + 1)}{\sqrt{\pi}2^{x+2}}.
\]

6 Problem 1.24

(a). Let the interval \([a, b]\) satisfy \( \int_a^b f(x)dx = 1 - \alpha \) and \( f(a) = f(b) > 0 \). Also let \( x^* \in [a, b] \) be the mode of \( f(x) \). We will show that for any interval \([a', b']\) such that \( b' - a' < b - a \), \( \int_{a'}^{b'} f(x)dx < 1 - \alpha \), thus proving that \([a, b]\) is the shortest interval. WLOG, assume \( a' \leq a \) and split the problem into two cases.

Case 1. Suppose \( b' \leq a \). Then \( a' \leq b' \leq a \leq x \) and

\[
\int_{a'}^{b'} f(x)dx \leq f(b')(b' - a') < f(a)(b - a) \leq \int_a^b f(x)dx = 1 - \alpha.
\]
Case 2. Otherwise, assume \( b' > a \). Then \( b' < b \) and
\[
\int_{a'}^{b'} f(x)dx = \int_{a}^{b} f(x)dx + \int_{a'}^{a} f(x)dx - \int_{b}^{b'} f(x)dx
\]
Hence we only need to show that \( \int_{a}^{a'} f(x)dx = \int_{b}^{b'} f(x)dx < 0 \) Note that \( a' \leq a \leq b' \leq b \), which implies \( \int_{a'}^{a} f(x)dx \leq f(a)(a - a') \) and \( \int_{b}^{b'} f(x)dx \geq f(b)(b - b') \). Hence
\[
\int_{a}^{a'} f(x)dx - \int_{b}^{b'} f(x)dx \leq f(a)(a - a') - f(b)(b - b')
\]
\[
= f(a)(a - a' - b + b')
\]
\[
= f(a)(b' - a') - (b - a) < 0.
\]
(b). If \( f \) is strictly monotone on either side of its mode, which we take to be 0, then \( f(x) = f(-x) \) for any \( x \in \mathbb{R} \). If \([a, b]\) is the shortest intercal such that \( \int_{a}^{b} f(x)dx = 1 - \alpha \), then
\[
f(a) = f(b) = f(-a) = f(-b) \text{ where } a < 0 < b.
\]
Now that \( f(a) = f(-b) \) for \( a, -b < 0 \) and \( f \) is strictly monotone, \( a = -b \) must hold.

7 Problem 1.28

(a). If \( X \sim G(\theta, \beta) \), then
\[
\pi(\theta|x, \beta) \propto \pi(\theta) \times (\beta x)^\theta / \Gamma(\theta)
\]
and a family of functions \( \pi(\theta) \) that are similar to the likelihood is given by
\[
\pi(\theta) \propto \xi\theta / \Gamma(\theta)^\alpha,
\]
where \( \xi > 0 \) and \( \alpha > 0 \) (in fact, \( \alpha \) could even be restricted to be an integer). This distribution is integrable when \( \alpha > 0 \) thanks to the Stirling approximation,
\[
\Gamma(\theta) \approx \theta^{\theta - 1/2} e^{-\theta}.
\]
(b). When \( X \sim Be(1, \theta), \theta \in \mathbb{N} \), we have
\[
f(x|\theta) = \frac{(1 - x)^{\theta-1}}{B(1, \theta)} = \frac{\Gamma(1 + \theta)(1 - x)^{\theta-1}}{\Gamma(\theta)} = \theta(1 - x)^{\theta-1}
\]
and this suggest using a gamma-like distribution on \( \theta \),
\[
\pi(\theta) \propto \theta^m e^{-\alpha\theta},
\]
where \( m \in \mathbb{N} \) and \( \alpha > 0 \). This function is clearly summable, due to the integrability of the gamma density, and conjugate.