THE UNIVERSITY OF NEW SOUTH WALES

DEPARTMENT OF STATISTICS

MATH5905 STATISTICAL INFERENCE

Part one: Decision theory. Bayes and minimax rules (SOLUTIONS)

Question 1: Please draw carefully the graph of the risk set before doing anything else.

a) d_3 since the minimal between the four values $\{6, 5, 3, 5\}$ is 3.
b) The rule d_3 again. Its minimax risk is 3.
c) The rule d_3 again. Its Bayes risk is equal to $\frac{1}{3} \times 2 + \frac{2}{3} \times 3 = 2\frac{2}{3}$.
d) Chooses d_2 and d_4 with probability 1/2 each.
e) All priors in the form $(p, 1-p)$ with $1 > p > 3/5$. Explanation: the slope $-\frac{p}{1-p}$ should be smaller than the slope $-\frac{3}{2}$ of $\overline{d_1 d_3}$.

Question 2: Since X is uniformly distributed in $[0, \theta)$ the density is $f(x, \theta) = \frac{1}{\theta} I_{[0,\theta)}(x)$ with $\mathbb{E}(X) = \frac{\theta}{2}$ and $\mathbb{E}(X^2) = \frac{\theta^2}{3}$. The rule is unbiased when $\mu = 2$ since $\mathbb{E}(2X) = 2 \cdot \frac{\theta}{2} = \theta$ holds. Now for any fixed value of μ we have

$$\mathbb{E}\big[(\theta - \mu X)^2\big] = \theta^2(1 - \mu + \mu^2/3).$$

When $\mu = \frac{3}{2}$ the latter mean squared error is equal to $\frac{\theta^2}{4}$. Now, we get

$$\mathbb{E}\big[(\theta - \mu X)^2\big] - \mathbb{E}\big[(\theta - \frac{3}{2}X)^2\big] = \frac{\mu^2\theta^2}{3} - \mu\theta^2 + \frac{3\theta^2}{4} = \frac{\theta^2}{12}(2\mu - 3)^2 \geq 0$$

the rule $\frac{3}{2}X$ will be uniformly better than any other rule in the form μX. That is, any rule in the form μX would be inadmissible unless $\mu = 3/2$.

Question 3: i) The likelihood times the prior gives

$$f(\mathbf{x}|\theta)\tau(\theta) = k\theta^n e^{-\theta(\sum_{i=1}^n x_i + k)}$$

and the marginal density of X is

$$g(\mathbf{x}) = \int_0^\infty f(\mathbf{x}|\theta)\tau(\theta)d\theta = k\int_0^\infty \theta^n e^{-\theta(\sum_{i=1}^n x_i + k)}d\theta$$

Now we change the variables: set

$$\theta(\sum_{i=1}^n x_i + k) = y, \qquad d\theta = \frac{dy}{(\sum_{i=1}^n x_i + k)}$$

and get:

$$g(\mathbf{x}) = \frac{k}{(\sum_{i=1}^n x_i + k)^{n+1}}\int_o^\infty y^n e^{-y}dy = \frac{k\Gamma(n+1)}{(\sum_{i=1}^n x_i + k)^{n+1}}$$

Hence

$$h(\theta|\mathbf{x}) = \frac{\theta^n e^{-\theta(\sum_{i=1}^n x_i + k)}}{\Gamma(n+1)(\frac{1}{\sum_{i=1}^n x_i + k})^{n+1}}, \quad \theta > 0.$$

Then by recalling the general definition of a Gamma(α, β) density:

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

we see that

$$h(\theta|\mathbf{x}) \sim \text{ Gamma}(n+1, \frac{1}{\sum_{i=1}^{n} x_i + k}).$$

Note: we did note really have to determine the normalizing constant as we did above. There is an easier approach based on looking at the joint density

$$f(\mathbf{x}|\theta)\tau(\theta) = k\theta^n e^{-\theta(\sum_{i=1}^{n} x_i + k)}$$

we can identify that (up to a normalizing constant) that this is a

$$Gamma\Big(n+1, \frac{1}{\sum_{i=1}^{n} x_i + k}\Big)$$

density and hence the posterior $h(\theta|\mathbf{x})$ has to be Gamma$(n+1, \frac{1}{\sum_{i=1}^{n} x_i + k})$.

ii) For a Bayes estimator with respect to quadratic loss, we have $\hat{\theta} = E(\theta|\mathbf{X})$, and for a Gamma$(\alpha, \beta)$ density it is known that the expected value is equal to $\alpha\beta$ hence we get immediately

$$\hat{\theta} = \frac{n+1}{\sum_{i=1}^{n} x_i + k}.$$

We could also calculate this directly:

$$\hat{\theta} = \int_0^\infty \theta h(\theta|\mathbf{x}) d\theta = \frac{(\sum_{i=1}^{n} x_i + k)^{n+1}}{\Gamma(n+1)} \int_0^\infty \theta^{n+1} e^{-\theta(\sum_{i=1}^{n} x_i + k)} d\theta$$

and after changing variables:

$$\theta(\sum_{i=1}^{n} x_i + k) = y, d\theta = \frac{dy}{(\sum_{i=1}^{n} x_i + k)}$$

we can continue the evaluation:

$$\begin{aligned}
\hat{\theta} &= \frac{\int_0^\infty e^{-y} y^{n+1} dy}{\Gamma(n+1)(\sum_{i=1}^{n} x_i + k)} \\
&= \frac{\Gamma(n+2)}{\Gamma(n+1)(\sum_{i=1}^{n} x_i + k)} \\
&= \frac{n+1}{\sum_{i=1}^{n} x_i + k}
\end{aligned}$$

Question 4: We have a single observation X only with density

$$f(x|\theta) = \frac{1}{\theta} I_{(x,\infty)}(\theta)$$

which implies that

$$g(x) = \int_0^\infty f(x|\theta)\tau(\theta) d\theta = \int_x^\infty \frac{1}{\theta}\theta e^{-\theta} d\theta = e^{-x}, \quad x > 0.$$

Hence

$$h(\theta|x) = \frac{f(x|\theta)\tau(\theta)}{g(x)} = \begin{cases} e^{x-\theta} & \text{if} \quad \theta > x \\ 0 & \text{if} \quad 0 < \theta < x \end{cases}$$

i) With respect to quadratic loss: The Bayesian estimator $\delta_\tau(x)$ is given by:

$$\delta_\tau(x) = \int_x^\infty \theta h(\theta|x) d\theta = \int_x^\infty \theta e^{x-\theta} d\theta = e^x \int_x^\infty \theta e^{-\theta} d\theta = e^x(xe^{-x} + e^{-x}) = x + 1.$$

ii) With respect to absolute value loss: The Bayesian estimator m solves the equation:

$$\int_m^\infty e^{x-\theta} d\theta = \frac{1}{2}$$

and we get: $e^{x-m} = \frac{1}{2} \Longrightarrow m - x = \ln 2 \Longrightarrow m = x + \ln 2.$

iii) To find the Bayes rule for the loss function $L_\eta(\theta, a) = (\theta - a)(\eta - I(\theta - a < 0))$ we need to find the action a such that

$$\inf_{a \in \mathcal{A}} Q(X, a)$$

which is the same as minimizing the Bayesian risk. Now:

$$\begin{aligned}
Q(X, a) &= \int_\Theta L(\theta, a) h(\theta|X) \mathrm{d}\theta \\
&= \int_x^\infty (\theta - a)(\eta - I(\theta < a)) h(\theta|X) \mathrm{d}\theta \\
&= \int_x^\infty \eta\theta h(\theta|X) \mathrm{d}\theta - \int_x^\infty \theta I(\theta < a) h(\theta|X) \mathrm{d}\theta \\
&\quad - \int_x^\infty a\eta h(\theta|X) \mathrm{d}\theta + \int_x^\infty a I(\theta < a) h(\theta|X) \mathrm{d}\theta \\
&= \eta \int_x^\infty \theta h(\theta|X) \mathrm{d}\theta - \int_x^a \theta h(\theta|X) \mathrm{d}\theta - a\eta \int_x^\infty h(\theta|X) \mathrm{d}\theta + a \int_0^a h(\theta|X) \mathrm{d}\theta
\end{aligned}$$

Therefore

$$\frac{\partial Q(X, a)}{\partial a} = -ah(a|X) - \eta + \int_x^a h(\theta|X) \mathrm{d}\theta + ah(a|X) = 0$$

which leads to the solution:

$$\int_x^a h(\theta|X) \mathrm{d}\theta = \eta$$

More specifically:

$$\int_x^a e^{x-\theta} \mathrm{d}\theta = \Big[- e^{x-\theta} \Big]_x^a = -e^{x-a} - (-e^0) = 1 - e^{x-a} = \eta$$

and the solution is:

$$a = x + \ln(1 - \eta)$$

Question 5: Let $\mathbf{X} = (X_1, \ldots, X_n)$ are the random variables. Setting $\mu_0 = x_0$ for convenience of the notation, we can write:

$$h(\mu|\mathbf{X}{=}\mathbf{x}) \propto e^{-\frac{1}{2}\sum_{i=0}^n (x_i - \mu)^2} \propto e^{-\frac{n+1}{2}[\mu^2 - 2\mu \frac{\sum_{i=0}^n x_i}{n+1}]}$$

Then by completing the square with the expression that does not depend on μ:

$$h(\mu|\mathbf{X}{=}\mathbf{x}) \propto e^{-\frac{n+1}{2}[\mu - \frac{\sum_{i=0}^n x_i}{n+1}]^2}$$

which implies that $h(\mu|\mathbf{X}{=}\mathbf{x})$, (being a density), must be the density of

$$N(\frac{\sum_{i=0}^n x_i}{n+1}, \frac{1}{n+1}).$$

Hence, the Bayes estimator (being the posterior mean) would be

$$(\sum_{i=0}^n x_i)/(n+1) = (\mu_0 + \sum_{i=1}^n x_i)/(n+1) = \frac{1}{n+1}\mu_0 + \frac{n}{n+1}\bar{X},$$

that is, the Bayes estimator is a convex combination of the mean of the prior and of $\bar{X}$. In this combination, the weight of the prior information diminishes quickly when the sample size increases. The **same** estimator is obtained with respect to absolute value loss.

Question 6: i) Since $X \sim \mathrm{Bin}(5, \theta)$ we have:

$$P(X = 0|\theta) = (1 - \theta)^5,$$

which means that the posterior of θ given the sample is

$$h(\theta|X = 0) \propto (1 - \theta)^5 \theta(1 - \theta)^4 = \theta(1 - \theta)^9.$$

Hence

$$h(\theta|X=0) = 110\theta(1-\theta)^9.$$

where $\frac{\Gamma(12)}{\Gamma(10)\Gamma(2)} = \frac{11!}{9!1!} = 110$. Then we get for the posterior probability given the sample:

$$P(\theta \in \Theta_0|X=0) = \int_0^{0.2} 110\theta(1-\theta)^9 d\theta = .6779$$

and we **accept** H_0 since the above posterior probability is greater than 0.5.

ii) Now

$$P(X=1|\theta) = 5(1-\theta)^4\theta,$$

which implies that the posterior of θ given the sample is

$$h(\theta|X=1) \propto (1-\theta)^4\theta(1-\theta)^4\theta = (1-\theta)^8\theta^2.$$

Hence

$$h(\theta|X=1) = \frac{\Gamma(12)}{\Gamma(9)\Gamma(3)}(1-\theta)^8\theta^2 = 495\theta^2(1-\theta)^8.$$

Then we get for the posterior probability given the sample:

$$P(\theta \in \Theta_0|X=1) = \int_0^{0.2} 495\theta^2(1-\theta)^8 d\theta = .3826 < \frac{1}{2}.$$

and we **reject** H_0 since the above posterior probability is smaller than 0.5.