THE UNIVERSITY OF NEW SOUTH WALES

DEPARTMENT OF STATISTICS

Solutions to selected exercises for MATH5905, Statistical Inference

Part one: Decision theory. Bayes and minimax rules

Question 1 Answers: 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 Note that for X uniformly distributed in $[0, \theta)$ we have the density $f(x, \theta) = \frac{1}{\theta} I_{[0,\theta)}(x)$ and from here we have easily $E(X) = \frac{\theta}{2}, E(X^2) = \frac{\theta^2}{3}$. The rule is unbiased when $\mu = 2 : E(2X) = \theta$ holds.

For any fixed value of μ we have $E(\theta - \mu X)^2 = \theta^2(1 - \mu + \mu^2/3)$. When $\mu = \frac{3}{2}$ the latter mean squared error is equal to $\frac{\theta^2}{4}$. Now, we get

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

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)

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

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

Now we change the variables: set $\theta(\sum_{i=1}^n x_i + k) = y, d\theta = \frac{dy}{(\sum_{i=1}^n x_i + k)}$ 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

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

Recalling the general definition of a Gamma(α, β) density:

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

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

NOTE: We **did NOT REALLY HAVE** to determine that normalising constant the way I showed above. Here is an **EASIER APPROACH**. Indeed just by 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 normalising constant this is a $Gamma(n+1, \frac{1}{\sum_{i=1}^{n} x_i + k})$ density 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 (α, β) density it is known that the expected value is equal to $\alpha\beta$ hence we get immediately $\hat{\theta} = \frac{n+1}{\sum_{i=1}^{n} x_i + k}$.
Of course, we could also calculate 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:

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

Question 4 Note that we have a SINGLE observation X only. Now: $f(x|\theta) = \frac{1}{\theta} I_{(x,\infty)}(\theta)$ implies that

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

Hence

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

i) With respect to quadratic loss: The Bayesian esimator $\delta_\tau(x)$ 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: $e^{x-m} = \frac{1}{2} \Longrightarrow m - x = ln2 \Longrightarrow m = x + ln2$.

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

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

Of course this also means (by completing the square with expression that does not depend on μ)

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

which implies that $h(\mu|\mathbf{X{=}x})$, (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 6i) $X \sim Bin(5, \theta)$. We have:

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

which means that the posterior of θ given the sample is $\propto (1-\theta)^5\theta(1-\theta)^4 = \theta(1-\theta)^9$. Hence

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

(Note: $\frac{\Gamma(12)}{\Gamma(10)\Gamma(2)} = \frac{11!}{9!1!} = 110$.) 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 $> \frac{1}{2.}$

ii) Now

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

which implies that the posterior of θ given the sample is $\propto (1-\theta)^4\theta(1-\theta)^4\theta = (1-\theta)^8\theta^2$. 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 $< \frac{1}{2.}$