UNSW Sydney
Department of Statistics
Term 1, 2020
MATH5905 - Statistical Inference
Assignment 1

Instructions: This assignment must be submitted no later than the beginning of the lecture at 6 pm on Friday, 13th March 2020. Please, declare on the first page that the assignment is your own work, except where acknowledged. State also that you have read and understood the University Rules in respect to Student Academic Misconduct. You need to submit the PDF file on **Moodle Turnitin** by 6 pm on Friday 13th March 2020. You will also need to bring a printed copy of your PDF file to the lecture by 6 pm on Friday 13th March 2020 as well.
Maximal number of pages: 8 pages

Problem 1
Consider a decision problem with parameter space $\Theta = \{\theta_1, \theta_2\}$ and a set of non randomized decisions $D = \{d_i, 1 \leq i \leq 7\}$ with risk points $\{R(\theta_1, d_i), R(\theta_2, d_i)\}$ as follows:

i	1	2	3	4	5	6	7
$R(\theta_1, d_i)$	1	3	7	13	15	13	10
$R(\theta_2, d_i)$	15	11	11	5	11	17	14

a) Find the minimax rule(s) amongst the **non-randomized** rules in D.

b) Plot the risk set of all **randomized** rules $\mathcal{D}$ generated by the set of rules in D.

c) Find the risk point of the minimax rule in $\mathcal{D}$ and determine its minimax risk.

d) Define the minimax rule in the set $\mathcal{D}$ in terms of rules in D.

e) For which prior on $\{\theta_1, \theta_2\}$ is the minimax rule in the set $\mathcal{D}$ also a Bayes rule?

f) Determine the Bayes rule and the Bayes risk for the prior $\left(\frac{2}{3}, \frac{1}{3}\right)$ on $\{\theta_1, \theta_2\}$.

g) For a small positive $\epsilon = 2$, illustrate on the risk set the risk points of all rules which are ϵ-minimax.

Solution:

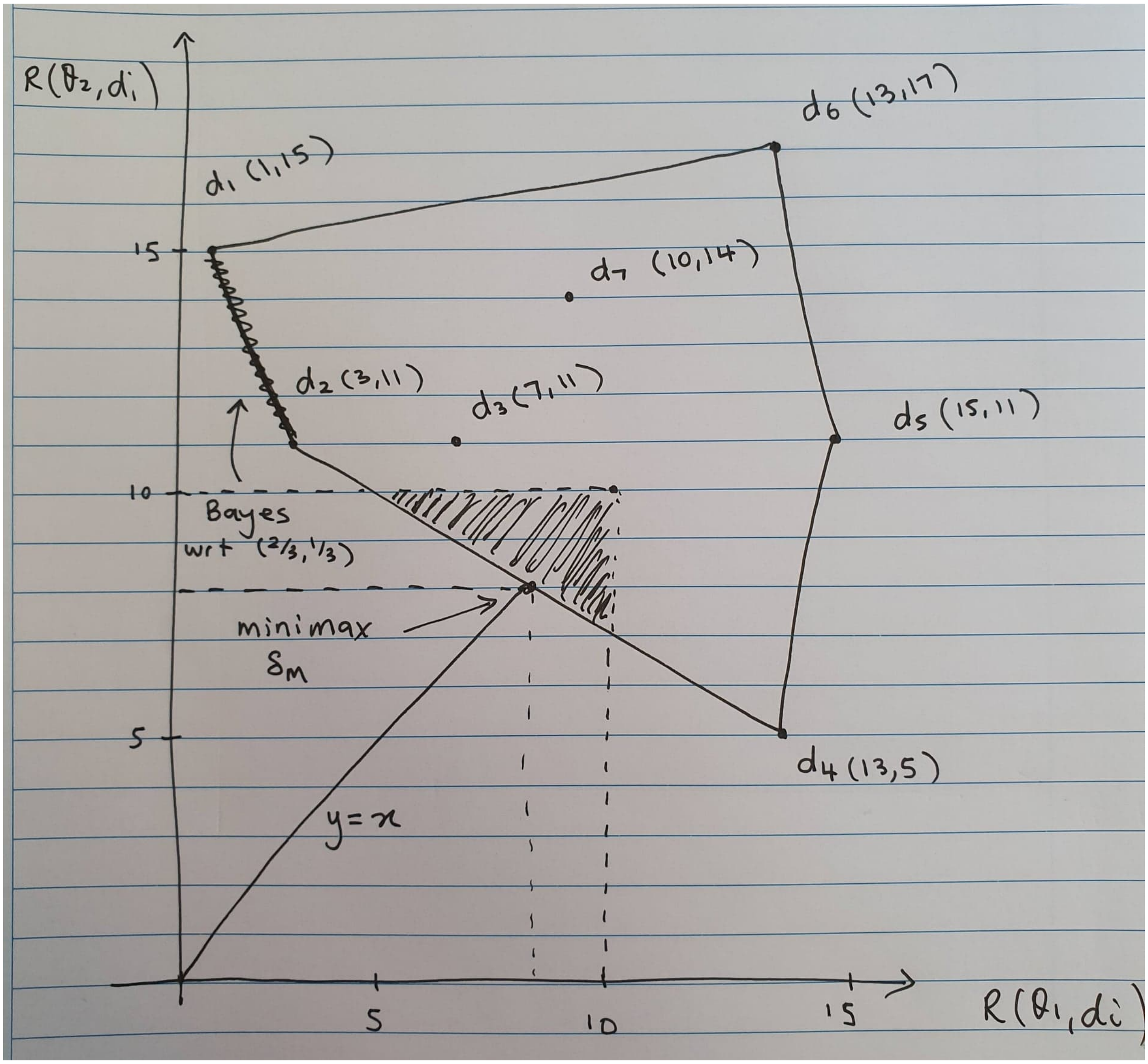
a) [1 mark] Find the minimax rule(s) amongst the **non-randomized** rules in D.

For each non-randomized decision rule we need to compute:

i	**1**	**2**	**3**	**4**	**5**	**6**	**7**
$\sup_{d \in D} R(\theta, d_i)$	**15**	**11**	**11**	**13**	**15**	**17**	**14**

Hence, $\inf_{\theta\in\Theta} \sup_{\theta\in\Theta} R(\theta, d)$ **is both** d_2 **and** d_3 **which are hence the minimax decisions in the set** D **with a minimax risk of** 11**.**

b) [2 marks] Plot the risk set of all **randomized** rules $\mathcal{D}$ generated by the set of rules in D.



c) [3 marks] Find the risk point of the minimax rule in $\mathcal{D}$ and determine its minimax risk.

Need to find the intersection of the lines $y = x$ and $\overline{d_2 d_4}$. That is,we want to solve

$$y = x \qquad \textbf{and} \qquad y - 5 = \frac{11-5}{3-13}(x - 13)$$

This is equivalent to

$$x - 5 = -\frac{6}{10}(x - 13).$$

Hence

$$x = y = \frac{10 \times 12.8}{16} = 8.$$

Hence the risk point δ_M of the minimax rule in $\mathcal{D}$ is $(8, 8)$ with a minimax risk of 8 which is less than the minimax risk in the set of non-randomised decision rules D with 11, as expected.

d) [2 marks] Define the minimax rule in the set $\mathcal{D}$ in terms of rules in D.

We try to express the rule δ_M in terms of d_2 and d_4 by finding a value $\alpha \in [0,1]$ such that

$$3\alpha + 13(1-\alpha) = 8 \qquad \textbf{and} \qquad 11\alpha + 5(1-\alpha) = 8.$$

Solving this leads to $\alpha = \frac{1}{2}$ and hence the randomized decision rule δ_M is to choose d_2 with probability 0.5 and choose d_4 with probability 0.5.

e) [2 marks] For which prior on $\{\theta_1, \theta_2\}$ is the minimax rule in the set $\mathcal{D}$ also a Bayes rule?

If the prior is $(p, 1-p)$ this leads to a line with a normal vector $(p, 1-p)$, that is, a slope with $\frac{-p}{1-p}$ and this slope should coincide with the slop of $\overline{d_2 d_4}$. Hence

$$\frac{-p}{1-p} = \frac{11-5}{3-13} = -\frac{3}{5}$$

should hold.

Solving this leads to $p = \frac{3}{8}$ and the least favourable prior with respect to which δ_M is Bayes is $(\frac{3}{8}, \frac{5}{8})$ on (θ_1, θ_2).

f) [3 marks] Determine the Bayes rule and the Bayes risk for the prior $\left(\frac{2}{3}, \frac{1}{3}\right)$ on $\{\theta_1, \theta_2\}$.

The line with normal vector $(\frac{2}{3}, 13)$ has slope $\frac{-2/3}{1/3} = -2$. When moving such a line "south-west" as much as possible but retaining intersection with the risk set, we end up with the line $\overline{d_1 d_2}$ as the set of decisions that correspond to the Bayes rule. This is because the slope of the line $\overline{d_1 d_2}$ is $\frac{15-11}{1-3} = -2$ and it is the most "south-west" on the risk-set.

It's Bayes risk is (just compute at either d_1 and d_2):

$$\frac{2}{3} \times 1 + \frac{1}{3} \times 15 = \frac{17}{3} \qquad \textbf{at } d_1$$

$$\frac{2}{3} \times 3 + \frac{1}{3} \times 11 = \frac{17}{3} \qquad \textbf{at } d_2.$$

g) [1 mark] For a small positive $\epsilon = 2$, illustrate on the risk set the risk points of all rules which are ϵ-minimax.

See shaded region on plot in part (b)

Problem 2

Suppose $X = (X_1, \ldots, X_n)$ are i.i.d. Poisson(θ) with density

$$f(x, \theta) = \frac{e^{-\theta}\theta^x}{x!}, \qquad x \in \{0, 1, 2, \ldots\}, \quad \theta > 0$$

and let θ have a Gamma(α, β) distribution with density

$$\tau(\theta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \theta^{\alpha-1} e^{-\theta/\beta}, \qquad \alpha, \beta > 0, \quad \theta > 0.$$

a) Find the posterior distribution for θ.

b) Hence or otherwise determine the Bayes estimator of θ with respect to the quadratic loss function $L(a, \theta) = (a - \theta)^2$.

c) Suppose the following nine observations were observed:

$$0, 1, 0, 0, 2, 3, 1, 1, 0.$$

Using a zero-one loss with the parameters $\alpha = 2$ and $\beta = 1$ for the prior, what is your decision when testing $H_0 : \theta \leq 1$ versus $H_1 : \theta > 1$. (You may use the `integrate` function in `R` or another numerical integration routine from your favourite programming package to answer the question.)

Solution:

a) [3 marks] Find the posterior distribution for θ.

We use the fact that the posterior is proportional to the prior and the likelihood:

$$\begin{aligned} p(\theta|X) &\propto L(X|\theta)\tau(\theta) \\ &= \frac{e^{-n\theta}\theta^{\sum_{i=1}^n X_i}}{\prod_{i=1}^n x_i} \frac{1}{\Gamma(\alpha)\beta^\alpha}\theta^{\alpha-1}e^{-\theta/\beta} \\ &\propto \theta^{\alpha+\sum_{i=1}^n X_i - 1}e^{-\theta(n+\frac{1}{\beta})}. \end{aligned}$$

Hence we can recognise this as a gamma density with parameters

$$\tilde{\alpha} = \alpha + \sum_{i=1}^n X_i \qquad \textbf{and} \qquad \tilde{\beta} = \frac{1}{n + \frac{1}{\beta}} = \frac{\beta}{n\beta + 1}$$

b) [1 mark] Hence or otherwise determine the Bayes estimator of θ with respect to the quadratic loss function $L(a, \theta) = (a - \theta)^2$.

The Bayes estimator with respect to quadratic loss is simply the posterior mean and hence

$$\hat{\theta}_{\mathbf{bayes}} = E(\theta|X) = \tilde{\alpha}\tilde{\beta} = (\alpha + \sum_{i=1}^n X_i)(\frac{\beta}{n\beta + 1}).$$

c) [4 marks] Suppose the following nine observations were observed:

$$0, 1, 0, 0, 2, 3, 1, 1, 0.$$

Using a zero-one loss with the parameters $\alpha = 2$ and $\beta = 1$ for the prior, what is your decision when testing $H_0 : \theta \leq 1$ versus $H_1 : \theta > 1$. (You may use the `integrate` function in `R` or another numerical integration routine from your favourite programming package to answer the question.)

The structure of the test is as follows:

$$\varphi = \begin{cases} 1 & \textbf{if} \quad P(\theta < 1|X) < 0.5 \\ 0 & \textbf{if} \quad P(\theta < 1|X) \geq 0.5 \end{cases}$$

Here we observe that $n = 9$, $\sum_{i=1}^9 x_i = 8$ **the posterior distribution is then:**

$$\theta|X \sim Gamma\left(2 + 8, \frac{1}{9 \times 1 + 1}\right) = Gamma(10, 0.1)$$

Then using R we can compute the posterior probability under these conditions as

$$P(\theta < 1|X) = \int_0^1 \theta \frac{1}{\Gamma(10)0.1^{10}} \theta^9 e^{-10\theta} \mathrm{d}\theta \approx 0.5420703$$

Hence, since the posterior probability is greater than 0.50 we do not reject H_0.

Problem 3
Determine the form of the Bayes decision rule in an estimation problem with a one-dimensional parameter $\theta \in \mathbb{R}^1$ and loss function

$$L\big(\theta, d\big) = \begin{cases} a\big(\theta - d\big) & \text{if} \quad d \le \theta, \\ b\big(d - \theta\big) & \text{if} \quad d > \theta, \end{cases}$$

where a and b are known positive constants.

Solution: [5 marks] We need to minimize the following

$$\begin{aligned} Q(X, d) &= \int_\theta L(\theta, d) h(\theta \mid X) \mathrm{d}\theta \\ &= \int_{-\infty}^d b(d-\theta) h(\theta \mid X) \mathrm{d}\theta + \int_d^\infty a(\theta - d) h(\theta \mid X) \mathrm{d}\theta \\ &= bd \int_{-\infty}^d h(\theta \mid X) \mathrm{d}\theta - b \int_{-\infty}^d \theta h(\theta \mid X) \mathrm{d}\theta + a \int_d^\infty \theta h(\theta \mid X) \mathrm{d}\theta - ad \int_d^\infty h(\theta \mid X) \mathrm{d}\theta \end{aligned}$$

Then by taking the derivative with respect to d we obtain

$$\begin{aligned} \frac{\partial}{\partial d} Q(X, d) &= b \int_{-\infty}^d h(\theta \mid X) \mathrm{d}\theta + bdh(d \mid X) - adh(d \mid X) - bdh(d \mid X) - a \int_d^\infty h(\theta \mid X) \mathrm{d}\theta + adh(d \mid X) \\ &= b \int_{-\infty}^d h(\theta \mid X) \mathrm{d}\theta - a \int_d^\infty h(\theta \mid X) \mathrm{d}\theta \end{aligned}$$

Then set this equal to zero

$$\begin{aligned} 0 &= b \int_{-\infty}^{d^*} h(\theta \mid X) \mathrm{d}\theta - a \int_{d^*}^\infty h(\theta \mid X) \mathrm{d}\theta \\ 0 &= bP(\Theta \le d^* \mid X = x) - aP(\Theta > d^* \mid X = x) \\ 0 &= bP(\Theta \le d^* \mid X = x) - a(1 - P(\Theta \le d^* \mid X = x)) \end{aligned}$$

For which we obtain

$$P(\Theta \le d^* \mid X = x) = \frac{a}{a+b}$$

which is simply the $\frac{a}{a+b}$ quantile of the posterior distribution. The indeed, delivers a minimum since

$$\frac{\partial^2}{\partial d^2} Q(X, d^*) = (b + a) h(d^* \mid X) > 0$$

since $a, b > 0$ and the posterior evaluated at d^* is also positive.

Problem 4
Suppose you are responsible for a deciding whether the light rail project in Sydney was to continue to be developed during a critial stage in its development. You must decide whether the project will continue or the project will be abandoned. The parameter $\theta \in (0, 1)$ measures the financial viability of the project and the project is deemed profitable when $\theta > 0.5$. You have some data x available to you that provides information about the parameter θ. If $\theta < 0.5$ then the cost to the taxpayer of continuing the project is $0.5 - \theta$ (in units of \$ million) whereas if $\theta > 0.5$ then the costs is zero since the project will be privatized if profitable. Furthermore, if $\theta > 0.5$ then the cost of abandoning the project is $\theta - 0.5$ due to the contractual arrangement for purchasing the machinery whereas if $\theta < 0.5$ then it is zero.

Therefore, two actions will be available to you:

$$a_0 : \text{continue the project} \qquad \text{and} \qquad a_1 : \text{abandon it.}$$

The prototype light-rail has been subjected to independent trials each with a probability θ of success. The data x then consists of the total number of trials required for the first successful result to be obtained. That is, one realization from a geometric distribution has been obtained. Two important parties are involved in the decision making. The Minister for Transport who has a prior density

$$\tau_1(\theta) = 30\theta^2(1-\theta)^2,$$

and the Prime Minister who has prior density

$$\tau_2(\theta) = 4\theta^3.$$

a) Derive the Bayesian decision rule in terms of the posterior mean of θ given x.

b) Determine the values of x for which there is the most serious ministerial disagreement between the two parties.

Solution:

a) [4 marks] Derive the Bayesian decision rule in terms of the posterior mean of θ given x.

We have two actions available to us: $a_0 =$ **continue or** $a_1 =$ **abandoned. The losses related to these two actions are**

$$L(\theta, a_0) = \begin{cases} \frac{1}{2} - \theta & \textbf{if} \quad \theta < 1/2 \\ 0 & \textbf{if} \quad \theta > 1/2 \end{cases},$$

$$L(\theta, a_0) = \begin{cases} 0 & \textbf{if} \quad \theta < 1/2 \\ \theta - \frac{1}{2} & \textbf{if} \quad \theta > 1/2 \end{cases},$$

For an optimal Bayes decision we need to compare:

$$Q(x, a_0) = \int_0^{1/2} (\frac{1}{2} - \theta) h(\theta|x) \mathrm{d}\theta = \frac{1}{2}\int_0^{1/2} h(\theta|x)\mathrm{d}\theta - \int_0^{1/2} \theta h(\theta|x)\mathrm{d}\theta$$

with

$$Q(x, a_1) = \int_{1/2}^1 (\theta - \frac{1}{2}) h(\theta|x)\mathrm{d}\theta = \int_{1/2}^1 \theta h(\theta|x)\mathrm{d}\theta - \frac{1}{2}\int_{1/2}^1 h(\theta|x)\mathrm{d}\theta.$$

Then a_0 **would be preferred to** a_1 **if** $Q(x, a_0) < Q(x, a_1)$**, or alternatively if** $Q(x, a_0) > Q(x, a_1)$ **then** a_1 **would be preferred - and there is hesitation when** $Q(x, a_0) = Q(x, a_1)$**.**

Now for the inequality $Q(x, a_0) < Q(x, a_1)$**:**

$$\frac{1}{2}\int_0^{1/2} h(\theta|x)\mathrm{d}\theta - \int_0^{1/2} \theta h(\theta|x)\mathrm{d}\theta < \int_{1/2}^1 \theta h(\theta|x)\mathrm{d}\theta - \frac{1}{2}\int_{1/2}^1 h(\theta|x)\mathrm{d}\theta$$

Observe that by adding $\pm\frac{1}{2}\int_0^{1/2} h(\theta|x)\mathrm{d}\theta$ **to the right hand side and noting that** $\int_0^1 h(\theta|x)\mathrm{d}\theta = 1$ **and rearranging we get:**

$$\frac{1}{2}\int_0^{1/2} h(\theta|x)\mathrm{d}\theta - \int_0^{1/2} \theta h(\theta|x)\mathrm{d}\theta < \int_{1/2}^1 \theta h(\theta|x)\mathrm{d}\theta - \frac{1}{2} + \frac{1}{2}\int_0^{1/2} h(\theta|x)\mathrm{d}\theta$$

Hence,

$$\frac{1}{2} < \int_0^1 \theta h(\theta|x)\mathrm{d}\theta$$

In other words, we choose $a_0 =$ continue, if $E(\theta|x) > 1/2$. Now for a Beta(α, β) distribution the expected value is $\frac{\alpha}{\alpha+\beta}$, which implies that in our case

$$E(\theta|x) = \frac{4}{x+6} \qquad \textbf{for the transport minister and}$$

$$E(\theta|x) = \frac{5}{x+5} \qquad \textbf{for the prime minister.}$$

b) [5 marks] For which values of x will there be most serious ministerial disagreement?

The observation scheme: we have $n = 1$ observation only from a <u>geometric</u> distribution with

$$f(x|\theta) = \theta(1-\theta)^{x-1}$$

where $\theta \in (0, 1)$ is the probability of success in a single trial until the first success. The two priors are

$$\tau_1(\theta) = 30\theta^2(1-\theta)^2 \qquad \textbf{of the transport ministers and}$$
$$\tau_2(\theta) = 4\theta^3 \qquad \textbf{of the prime minister.}$$

The two corresponding posteriors are:

$$h_1(\theta) \propto 30\theta^2(1-\theta)^2\theta(1-\theta)^{x-1} \propto \theta^3(1-\theta)^{x+1}$$
$$h_2(\theta) \propto 4\theta^3\theta(1-\theta)^{x-1} \propto \theta^4(1-\theta)^{x-1}$$

These can easily be identified as

$$h_1(\theta|x) \sim \textbf{Beta}(4, x+2)$$

$$h_2(\theta|x) \sim \textbf{Beta}(5, x)$$

From part a) we know that the transport minister wants the project to continue when $x = 1$, hesitates when $x = 2$ and wants the project to stop when $x = 3, 4, 5, \ldots$

On the other hand, the prime ministers wants to continue when $x = 1, 2, 3, 4$, hesitates when $x = 5$ and wants to stop when $x = 6, 7, 8, \ldots$

We have the most serious disagreement when $x = 3$ and $x = 4$ since the transport minister would like to abandoned the project while the prime minister would like it to continue.