UNSW Sydney
Department of Statistics
Term 1, 2024
MATH5905 - Statistical Inference
Assignment 1 Solutions

Problem One
A single die is tossed; then n coins are tossed where n is the number shown on the die. Show that the probability of exactly two heads is close to 0.2578.

Solution: Let $U1, U2, \ldots, U6$ denote the events "uppermost $1, 2, \ldots, 6$" respectively and A denote the event in question. We have $P(U1) = \cdots = P(U6) = \frac{1}{6}$. The formula of total probability gives

$$P(A) = P(A|U1) * P(U1) + P(A|U2)P(U2) + \ldots P(A|U6)P(U6) = \frac{1}{6}(P(A|U1) + \ldots P(A|U6)).$$

Obviously $P(A|U1) = 0$ holds. For the remaining conditional probabilities, considering 'Head' as a success and 'Tail' as a failure, we need to get the probability of a number of two successes out of i independent Bernoulli trials by calculating the ratio $\frac{\text{favourable outcomes}}{\text{total number of outcomes}}$. The total number of outcomes is 2^i and the favourable ones are $\binom{i}{2}$, that is, $P(A|Ui) = \frac{\binom{i}{2}}{2^i}, i = 2, 3, 4, 5, 6$. Hence we get

$$P(A) = \frac{1}{6}(0 + 1/4 + 3/8 + 6/16 + 20/64 + 30/128) = 33/128 \approx 0.2578.$$

Problem Two
A certain river floods every year. Suppose that the low-water mark is set at 1 and the high-water mark X has a distribution function

$$F_X(x) = P(X \leq x) = 1 - \frac{1}{x^2}, 1 \leq x < \infty$$

1. Verify that $F_X(x)$ is a cumulative distribution function
2. Find the density $f_X(x)$ (specify it on the whole real axis)
3. If the (same) low-water mark is reset at 0 and we use a unit of measurement that is $\frac{1}{10}$ of that used previously, express the random variable Z for the new measurement as a function of X. Then find precisely the cumulative distribution function and the density of Z.

Solution: a) Obviously $\lim_{x\to-\infty} F_X(x) = 0$ (as $F_X(x)$ is a constant 0 for $x \leq 1$). Also,

$$\lim_{x\to\infty} F_X(x) = \lim_{x\to\infty} 1 - \frac{1}{x^2} = 1.$$

For $x > 1$, $\frac{d}{dx}F_X(x) = 2/x^3 > 0$ which implies that $F_X(x)$ is increasing. So-for all x on the real axis, $F_X(x)$ is non-decreasing.

b) The density as a derivative was calculated in a) already for $x > 1$ and it is zero for $x \leq 1$.

c) $F_Z(z) = P(Z \leq z) = P(10(X - 1) \leq z) = P(X \leq (z/10) + 1) = F_X((z/10) + 1)$. Hence

$$F_Z(z) = 1 - (\frac{1}{[z/10 + 1]^2})$$

when $z > 0$ (and zero else).

The density is obtained via differentiation of the cdf leading to

$$f_Z(z) = \frac{1}{5(z/10 + 1)^3}, z > 0$$

(and zero else).

Problem 3
Suppose $X = (X_1, \dots, X_n)$ are i.i.d. Poisson(θ) with a probability mass function

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

The prior on the unknown parameter θ is assumed to be 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 ten observations were observed:

$$4, 0, 1, 0, 0, 2, 3, 1, 1, 0.$$

Using a zero-one loss with the parameters $\alpha = 2$ and $\beta = 1$ for the prior, what is your decision when testing $H_0 : \theta \le 1.2$ versus $H_1 : \theta > 1.2$. (You may use the `integrate` function in `R` or another numerical integration routine from your favourite programming package to answer the question.)

Solution:

a) 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 recognise this as a gamma density with parameters

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

b) From lectures, the Bayes estimator with respect to quadratic loss is the posterior mean given the sample.

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

c) The structure of the test is as follows:

$$\varphi = \begin{cases} 1 & \text{if} \quad P(\theta < 1.2|X) < 0.5 \\ 0 & \text{if} \quad P(\theta < 1.2|X) \ge 0.5 \end{cases}$$

We have $n = 10$, $\sum_{i=1}^{10} x_i = 12$ so the posterior distribution given the sample is:

$$\theta|X \sim Gamma\left(2+12, \frac{1}{10\times 1+1}\right) = Gamma(14, 1/11)$$

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

$$P(\theta < 1.2|X) = \int_0^{1.2} \frac{1}{\Gamma(14)(1/11)^{14}} \theta^{13} e^{-11\theta} \mathrm{d}\theta \approx 0.44893$$

Hence, since the posterior probability is smaller than 0.50 we reject H_0.

Problem Four

The Premier of NSW has to take an important decision whether or not to include an additional line (extension) to Sydney's metro system. The decision is based on the financial viability of the extension. He is willing to apply decision theory in making his decision. He uses the independent opinion of two consulting experts. The data he uses is the number X of viability recommendations of the two experts (so $X = 0, 1$ or 2).

If the Premier decides not to go ahead and the extension turns out not to be financially viable or if he decides to go ahead and the project is financially viable, nothing is lost. If the project is not financially viable and he has decided to go ahead, his subjective judgement is that his loss would equal three times the loss of not going ahead but the project is financially viable.

The Premier has investigated the history of viability predictions of the two consulting experts and it is as follows. When a project is financially viable, both experts have correctly predicted its viability with probability 4/5 (and wrongly with a probability 1/5). When a project has not been financially viable, both experts had a prediction of 3/5 for it to be viable. The Premier listens to the recommendations of the two experts and makes his decision based on the value of X.

a) There are two possible actions in the action space $\mathcal{A} = \{a_0, a_1\}$ where action a_0 is to go ahead and action a_1 is not to go ahead with the extension. There are two states of nature $\Theta = \{\theta_0, \theta_1\}$ where $\theta_0 = 0$ represents "Extension is financially viable" and $\theta_1 = 1$ represents "Extension is not financially viable". Define the appropriate loss function $L(\theta, a)$ for this problem.

b) Compute the probability mass function (pmf) for X under both states of nature.

c) The complete list of all the non-randomized decisions rules D based on x is given by:

	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8
$x = 0$	a_0	a_1	a_0	a_1	a_0	a_1	a_0	a_1
$x = 1$	a_0	a_0	a_1	a_1	a_0	a_0	a_1	a_1
$x = 2$	a_0	a_0	a_0	a_0	a_1	a_1	a_1	a_1

For the set of non-randomized decision rules D compute the corresponding risk points.

d) Find the minimax rule(s) among the **non-randomized** rules in D.

e) Sketch the risk set of all **randomized** rules $\mathcal{D}$ generated by the set of rules in D. You might want to use `R` (or your favorite programming language) to make this sketch more precise.

f) Suppose there are two decisions rules d and d'. The decision d strictly dominates d' if $R(\theta, d) \leq R(\theta, d')$ for all values of θ and $R(\theta, d) < (\theta, d')$ for at least one value θ. Hence, given a choice between d and d' we would always prefer to use d. Any decision rules which is strictly dominated by another decisions rule (as d' is in the above) is said to be inadmissible. Correspondingly, if a decision rule d is not strictly dominated by any other decision rule then it is admissible. Show on the risk plot the set of randomized decisions rules that correspond to the admissible decision rules.

g) Find the risk point of the minimax rule in the set of randomized decision rules $\mathcal{D}$ and determine its minimax risk. Compare the two minimax risks of the minimax decision rule in D and in $\mathcal{D}$. Comment.

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

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

j) Prior to listening to the two experts, the Premier's belief in the viability is 50%. Find the Bayes rule and the Bayes risk with respect to his prior.

k) For a small positive $\epsilon = 0.1$, illustrate on the risk set the risk points of all rules which are ϵ-minimax.

Solution: a) There are two actions: a_0 : decide to go ahead and a_1 : do not go ahead. Two states of nature, $\theta_0 = 0$, "Extension is financially viable", and $\theta_1 = 1$ "Extension is not financially viable". Let x be the number of experts predicting financial viability. The loss function is given by

$$L(\theta_1, a_0) = 3, L(\theta_1, a_1) = 0, L(\theta_0, a_0) = 0, L(\theta_0, a_1) = 1.$$

b) The pmf for both states of nature:

x	$p(x\|\theta = 0)$	$p(x\|\theta = 1)$
0	$\frac{1}{5} \cdot \frac{1}{5} = \frac{1}{25}$	$\frac{2}{5} \cdot \frac{2}{5} = \frac{4}{25}$
1	$2 \cdot \frac{1}{5} \cdot \frac{4}{5} = \frac{8}{25}$	$2 \cdot \frac{3}{5} \cdot \frac{2}{5} = \frac{12}{25}$
2	$\frac{4}{5} \cdot \frac{4}{5} = \frac{16}{25}$	$\frac{3}{5} \cdot \frac{3}{5} = \frac{9}{25}$

c)
There are $2^3 = 8$ non-randomized decision rules:

	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8
$x = 0$	a_0	a_1	a_0	a_1	a_0	a_1	a_0	a_1
$x = 1$	a_0	a_0	a_1	a_1	a_0	a_0	a_1	a_1
$x = 2$	a_0	a_0	a_0	a_0	a_1	a_1	a_1	a_1

Calculation of the risk points $\{R(\theta_0, d_i), R(\theta_1, d_i)\}$ is as follows:
For $d_1 = (0, 3)$:

$$R(\theta_0, d_1) = L(\theta_0, a_0)P(x = 0|\theta_0) + L(\theta_0, a_0)P(x = 1|\theta_0) + L(\theta_0, a_0)P(x = 2|\theta_0) = 0$$

as $L(\theta_0, a_0) = 0$.

$$R(\theta_1, d_1) = L(\theta_1, a_0)P(x = 0|\theta_1) + L(\theta_1, a_0)P(x = 1|\theta_1) + L(\theta_1, a_0)P(x = 2|\theta_1) = 3$$

as $L(\theta_1, a_0) = 3$.
For $d_2 = (1/25, 63/25)$:

$$\begin{aligned} R(\theta_0, d_2) &= L(\theta_0, a_1)P(x = 0|\theta_0) + L(\theta_0, a_0)P(x = 1|\theta_0) + L(\theta_0, a_0)P(x = 2|\theta_0) \\ &= 1 \times \frac{1}{25} + 0 \times \frac{8}{25} + 0 \times \frac{16}{25} \\ &= \frac{1}{25} \end{aligned}$$

$$\begin{aligned} R(\theta_1, d_2) &= L(\theta_1, a_1)P(x = 0|\theta_1) + L(\theta_1, a_0)P(x = 1|\theta_1) + L(\theta_1, a_0)P(x = 2|\theta_1) \\ &= 0 + 3 \times \frac{12}{25} + 3 \times \frac{9}{25} \\ &= \frac{63}{25} \end{aligned}$$

For $d_3 = (8/25, 39/25)$**:**

$$\begin{aligned}
R(\theta_0, d_3) &= L(\theta_0, a_0)P(x = 0|\theta_0) + L(\theta_0, a_1)P(x = 1|\theta_0) + L(\theta_0, a_0)P(x = 2|\theta_0) \\
&= 0 \times \frac{1}{25} + 1 \times \frac{8}{25} + 0 \times \frac{16}{25} \\
&= \frac{8}{25}
\end{aligned}$$

$$\begin{aligned}
R(\theta_1, d_3) &= L(\theta_1, a_0)P(x = 0|\theta_1) + L(\theta_1, a_1)P(x = 1|\theta_1) + L(\theta_1, a_0)P(x = 2|\theta_1) \\
&= 3 \times \frac{4}{25} + 0 \times \frac{12}{25} + 3 \times \frac{9}{25} \\
&= \frac{39}{25}
\end{aligned}$$

For $d_4 = (9/25, 27/25)$**:**

$$\begin{aligned}
R(\theta_0, d_4) &= L(\theta_0, a_1)P(x = 0|\theta_0) + L(\theta_0, a_1)P(x = 1|\theta_0) + L(\theta_0, a_0)P(x = 2|\theta_0) \\
&= 1 \times \frac{1}{25} + 1 \times \frac{8}{25} + 0 \times \frac{16}{25} \\
&= \frac{9}{25}
\end{aligned}$$

$$\begin{aligned}
R(\theta_1, d_4) &= L(\theta_1, a_1)P(x = 0|\theta_1) + L(\theta_1, a_1)P(x = 1|\theta_1) + L(\theta_1, a_0)P(x = 2|\theta_1) \\
&= 0 \times \frac{4}{25} + 0 \times \frac{12}{25} + 3 \times \frac{9}{25} \\
&= \frac{27}{25}
\end{aligned}$$

For $d_5 = (16/25, 48/25)$**:**

$$\begin{aligned}
R(\theta_0, d_5) &= L(\theta_0, a_0)P(x = 0|\theta_0) + L(\theta_0, a_0)P(x = 1|\theta_0) + L(\theta_0, a_1)P(x = 2|\theta_0) \\
&= 0 \times \frac{1}{25} + 0 \times \frac{8}{25} + 1 \times \frac{16}{25} \\
&= \frac{16}{25}
\end{aligned}$$

$$\begin{aligned}
R(\theta_1, d_5) &= L(\theta_1, a_0)P(x = 0|\theta_1) + L(\theta_1, a_0)P(x = 1|\theta_1) + L(\theta_1, a_1)P(x = 2|\theta_1) \\
&= 3 \times \frac{4}{25} + 3 \times \frac{12}{25} + 0 \times \frac{9}{25} \\
&= \frac{48}{25}
\end{aligned}$$

For $d_6 = (17/25, 36/25)$**:**

$$\begin{aligned}
R(\theta_0, d_6) &= L(\theta_0, a_1)P(x = 0|\theta_0) + L(\theta_0, a_0)P(x = 1|\theta_0) + L(\theta_0, a_1)P(x = 2|\theta_0) \\
&= 1 \times \frac{1}{25} + 0 \times \frac{8}{25} + 1 \times \frac{16}{25} \\
&= \frac{17}{25}
\end{aligned}$$

$$
\begin{aligned}
R(\theta_1, d_6) &= L(\theta_1, a_1)P(x=0|\theta_1) + L(\theta_1, a_0)P(x=1|\theta_1) + L(\theta_1, a_1)P(x=2|\theta_1) \\
&= 0 \times \frac{4}{25} + 3 \times \frac{12}{25} + 0 \times \frac{9}{25} \\
&= \frac{36}{25}
\end{aligned}
$$

For $d_7 = (24/25, 12/25)$**:**

$$
\begin{aligned}
R(\theta_0, d_7) &= L(\theta_0, a_0)P(x=0|\theta_0) + L(\theta_0, a_1)P(x=1|\theta_0) + L(\theta_0, a_1)P(x=2|\theta_0) \\
&= 0 \times \frac{1}{25} + 1 \times \frac{8}{25} + 1 \times \frac{16}{25} \\
&= \frac{24}{25}
\end{aligned}
$$

$$
\begin{aligned}
R(\theta_1, d_7) &= L(\theta_1, a_0)P(x=0|\theta_1) + L(\theta_1, a_1)P(x=1|\theta_1) + L(\theta_1, a_1)P(x=2|\theta_1) \\
&= 3 \times \frac{4}{25} + 0 \times \frac{12}{25} + 0 \times \frac{9}{25} \\
&= \frac{12}{25}
\end{aligned}
$$

For $d_8 = (0, 1)$**:**

$$R(\theta_0, d_8) = L(\theta_0, a_1)P(x=0|\theta_0) + L(\theta_0, a_1)P(x=1|\theta_0) + L(\theta_0, a_1)P(x=2|\theta_0) = 1$$

as $L(\theta_0, a_1) = 1$.

$$R(\theta_1, d_8) = L(\theta_1, a_1)P(x=0|\theta_1) + L(\theta_1, a_1)P(x=1|\theta_1) + L(\theta_1, a_1)P(x=2|\theta_1) = 0$$

as $L(\theta_1, a_1) = 0$.
This leads to the following risk points:

	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8
$R(\theta_0, d_i)$	0	$\frac{1}{25}$	$\frac{8}{25}$	$\frac{9}{25}$	$\frac{16}{25}$	$\frac{17}{25}$	$\frac{24}{25}$	1
$R(\theta_1, d_i)$	3	$\frac{63}{25}$	$\frac{39}{25}$	$\frac{27}{25}$	$\frac{48}{25}$	$\frac{36}{25}$	$\frac{12}{25}$	0

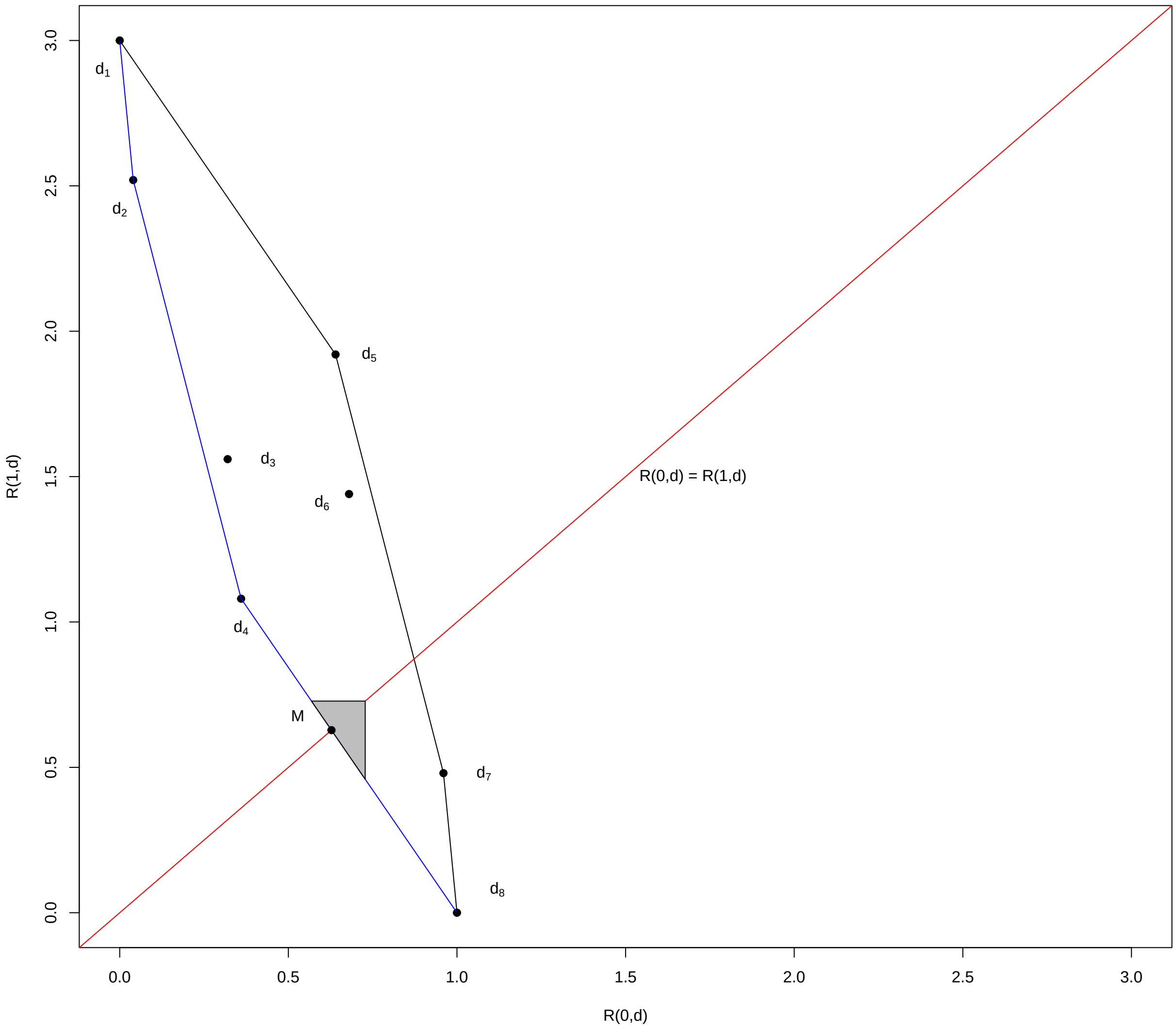
d) For each non-randomized decision rule we need to compute:

	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8
$\sup_{\theta \in \{\theta_0, \theta_1\}} R(\theta, d_i)$	3	$\frac{63}{25}$	$\frac{39}{25}$	$\frac{27}{25}$	$\frac{48}{25}$	$\frac{36}{25}$	$\frac{24}{25}$	1

Hence, $\inf_{d \in D} \sup_{\theta \in \Theta} R(\theta, d)$ is d_7 which therefore the minimax decision in the set D with a minimax risk of 24/25.

e) Sketch of the randomized rules $\mathcal{D}$ generated by the set of non-randomized decision rules D: see attached graph of the set.

f) The admissible rules are those on the "south-west boundary" of the risk set: any convex combination of d_8 and d_4, or d_4 and d_2, or d_2 and d_1.The randomized decisions rules that correspond to admissible rules



d1
d2
d3
d4
d5
d6
d7
d8
M
R(0,d) = R(1,d)
R(1,d)
R(0,d)
0.0
0.5
1.0
1.5
2.0
2.5
3.0

are colored in blue on the graph.

g) The minimax decision rule in the set is given by the intersection of the lines $y = x$ and the risk set towards the south-most boundary. The line $\overline{d_8 d_4}$ has an equation

$$y = \frac{\frac{27}{25} - 0}{\frac{9}{25} - 1}(x - 1) = -\frac{27}{16}x + \frac{27}{16}$$

To find the intersection, we solve

$$x = -\frac{27}{16}x + \frac{27}{16}$$

which gives $x = y = \frac{27}{43} = 0.627907$ as the coordinates of the intersection point M. Hence the risk point of the minimax rule δ^* in $\mathcal{D}$ is $\delta^* = (0.627907, 0.627907)$ with a minimax risk of 0.627907. We also realize that the risk of the minimax rule in the set D is reduced further in the larger set of randomized decision rules $\mathcal{D}$ (as $24/25 = 0.96 > 0.627907$).

h) To express the rule M in terms of d_4 and d_8 we need to find $\alpha \in [0, 1]$ such that

$$27/43 = 9/25\alpha + 1 * (1 - \alpha) \text{ and } 27/43 = 27/25\alpha + 0 * (1 - \alpha).$$

Each of these two relations gives the same solution $\alpha = \frac{25}{43} = 0.5814$. Therefore the randomized minimax rule chooses d_4 with probability $\alpha = 0.5814$ and d_8 with probability 0.4186.

i) Suppose 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 slope of $\overline{d_4 d_8}$. Hence

$$\frac{-p}{1 - p} = \frac{-27}{16}$$

should hold.

Solving this leads to $p = \frac{27}{43}$ and the least favourable prior with respect to which δ_M is Bayes is $(0.627907, 0.372093)$ on (θ_0, θ_1). As we know from the lectures, with respect to this prior, the minimax rule M is also a Bayes rule.

j) The line with normal vector $(\frac{1}{2}, \frac{1}{2})$ has slope -1. When moving such a line "south-west" as much as possible but retaining intersection with the risk set, we end up with the intersection point d_8. Hence d_8 is the Bayes decision that corresponds to the manager's prior.

It's Bayes risk is:

$$\frac{1}{2} \times 1 + \frac{1}{2}0 = 1/2$$

k) The shaded area on the graph. From the point $(27/43+0.1, 27/43+0.1)$, put a vertical and a horizontal line. The region in the risk set that is southwest of the quadrant defined by these lines is the relevant one.

Problem Five

Let a continuous random variable T be the length of live of an electrical component. The hazard function $h_T(t)$ associated with T is defined as

$$h_T(t) = \lim_{\eta \to 0} \frac{P(t \le T < t + \eta | T \ge t)}{\eta}.$$

(In other words, $h_T(t)$ describes the rate of change of the probability that the component survives a little past time t given that the component survives to time t.

a) Denoting by $F_T(t)$ and $f_T(t)$ the cdf and the density of T respectively, show that

$$h_T(t) = \frac{f_T(t)}{1 - F_T(t)} = -\frac{d}{dt}\log(1 - F_T(t).$$

b) Verify that if T is exponentially distributed, i.e.,

$$f_T(t) = \beta e^{-t\beta}, t > 0$$

where $\beta > 0$ is a parameter of the distribution then $h_T(t) = \beta$, that is, the hazard function is a constant for the exponential distribution.

c) Verify that if T is logistic with parameters μ (real) and $\beta > 0$, i.e.,

$$F_T(t) = \frac{1}{1 + \exp\left(-\frac{t-\mu}{\beta}\right)}$$

then $h_T(t) = \frac{1}{\beta} F_T(t)$.

Solution: a) We have

$$P(t \le T \le t + \eta | t \le T) = \frac{P(t \le T \le t + \eta)}{P(t \le T)} = \frac{F_T(t+\eta) - F_T(t)}{1 - F_T(t)}.$$

Therefore

$$h_T(t) = \frac{1}{1 - F_T(t)} \lim_{\eta \to 0} \frac{F_T(t+\eta) - F_T(t)}{\eta} = \frac{F_T'(t)}{1 - F_T(t)} = \frac{f_T(t)}{1 - F_T(t)}.$$

We also see directly that

$$-\frac{d}{dt}(\log[1 - F_T(t)]) = -\frac{f_T(t)}{1 - F_T(t)} = h_T(t)$$

holds.

b) Since $f_T(t) = \beta e^{-\beta t}, F_T(t) = 1 - e^{-\beta t}$ per definition, direct substitution in the formula for the hazard from a) gives

$$h_T(t) = \frac{\beta e^{-\beta t}}{1 - (1 - e^{-\beta t})} = \beta.$$

c) Taking the derivative of $F_T(t)$ we get

$$f_T(t) = \frac{1}{\beta} \frac{e^{-(t-\mu)/\beta}}{(1 + e^{-(t-\mu)/\beta})^2}.$$

Direct substitution in the formula for the hazard leads then to

$$h_T(t) = \frac{1}{\beta} \frac{e^{-(t-\mu)/\beta}}{(1 + e^{-(t-\mu)/\beta})^2} \frac{1 + e^{-(t-\mu)/\beta}}{e^{-(t-\mu)/\beta}} = \frac{1}{\beta} F_T(t).$$