Sheet 3

Problem 1

For each decision problem we need to find the Samplespace Ω, the decision space D, the loss function L and the parameter θ for the distribution.

One Week case

Because the student can either have normal heating or additional heating, the decision space has two elements.

$$D := \{\text{normal heating, additional heating}\}$$

The state of the loss is decided by the state of the student:

$$\Omega = \{\text{healthy, sick}\}$$

And the probability of sickness is given by $\theta \in [0, 1]$ with the state of the Student being a r.v. X:

$$X : \Omega \to \{0, 1\} \quad X \sim Bernoulli(\theta)$$

Then we can look at the costs depending on the action and the state of the student for the first week.

state/action	normal heating	additional heating
healthy	0€	6€
sick	20€	26€

From there we can define the loss function under normal heating:

$$L(\theta|\text{normal heating}) = \begin{cases} 0€ \text{ if } \theta = \text{healthy} \\ 20€ \text{ if } \theta = \text{sick} \end{cases}$$

…and under additional heating:

$$L(\theta|\text{additional heating}) = \begin{cases} 6€ \text{ if } \theta = \text{healthy} \\ 26€ \text{ if } \theta = \text{sick} \end{cases}$$

We also know the posterior probabilities:

$$\begin{aligned} P(X = \text{ healthy}|\delta = \text{normal heating}) = 0.7 \\ P(X = \text{ sick}|\delta = \text{normal heating}) = 0.3 \\ P(X = \text{ healthy}|\delta = \text{additional heating}) = 0.9 \\ P(X = \text{ sick}|\delta = \text{additional heating}) = 0.1 \end{aligned}$$

The student will choose the action, that yields the lowest risk:

$$\begin{aligned}R(\theta|\text{ normal heating}) &= \mathbb{E}(L(\theta|\text{normal heating}) \\ &= P(\theta = \text{healthy}|\delta = \text{normal heating}) \cdot L(\text{healthy}|\text{normal heating}) \\ &+ P(\theta = \text{sick}|\delta = \text{normal heating}) \cdot L(\text{sick}|\text{normal heating}) \\ &= 0.7 \cdot 0 + 0.3 \cdot 20 = 6€\end{aligned}$$

Applying the same Formula to compute $R(\theta|\text{additional heating})$:

$$R(\theta|\text{additional heating}) = 0.9 \cdot 6€ + 0.1 \cdot 26€ = 5.4€ + 2.6€ = 8€$$

Therefore the student will decide to use normal heating.

Two week case

In this case we have to distinguish between the states sick in the first week and second week:

$$\Omega = \{\text{healthy}, \text{sick first week}, \text{sick second week}\}$$

Because the student sticks to a decision throughout the cold period, we have still two decisions in the decision space:

$$D := \{\text{normal heating}, \text{additional heating}\}$$

Weekly heating costs for additional heating are $6€$. Deciding for additional heating yields costs independent of the state of $12€$.

Applying these states yields the new loss function:

$$\begin{aligned}L(\theta|\text{normal heating}) &:= \begin{cases} 0€ \text{ if } \theta = \text{healthy} \\ 20€ \text{ if } \theta = \text{sick second week} \\ 40€ \text{ if } \theta = \text{sick first week}\end{cases} \\ L(\theta|\text{additional heating}) &:= \begin{cases} 12€ \text{ if } \theta = \text{healthy} \\ 32€ \text{ if } \theta = \text{sick second week} \\ 52€ \text{ if } \theta = \text{sick first week}\end{cases}\end{aligned}$$

Given that the student gets sick during the first week, he/ she will stay sick with prob. 100%.:

$$\begin{aligned}P(X = \text{sick first week}|\delta = \text{normal heating}) &= 0.3 \cdot 1 = 0.3 \\ P(X = \text{sick second week}|\delta = \text{normal heating}) &= P(X = \text{healthy}|\delta = \text{normal heating} \\ &\cdot P(X = \text{sick}|\delta = \text{normal heating}) \\ &= 0.7 \cdot 0.3 = 0,21 \\ P(X = \text{healthy}|\delta = \text{normal heating}) &= 0.7^2 = 0.49\end{aligned}$$

Similarly we can compute the probabilities for the states under $\delta = \text{additional heating}$:

$$P(X = \text{sick first week}|\delta = \text{additional heating}) = 0.1$$
$$P(X = \text{sick second week}|\delta = \text{additional heating}) = 0.1 \cdot 0.9 = 0.09$$
$$P(X = \text{healthy}|\delta = \text{additional heating}) = 0.9^2 = 0.81$$

Therefore we can compute the risk:

$$\begin{aligned} R(\theta|\delta = \text{normal heating}) &= \mathbb{E}(L(\theta|\text{normal heating}) \\ &= P(\theta = \text{sick first week}|\delta = \text{normal heating}) \\ &\cdot L(\text{sick first week}|\delta = \text{normal heating}) \cdots \\ &= 0.3 \cdot 40€ + 0.21 \cdot 20€ + 0.49 \cdot 0€ = 16,2€ \\ R(\theta|\delta = \text{additional heating}) &= 0.1 \cdot 52€ + 0.09 \cdot 32€ + 0.81 \cdot 12€ = 17,8€ \end{aligned}$$

The right decision riskwise would therefore still be **normal heating.**

Three week case

We get one more additional case **sick in third week:**

$$\Omega = \{\text{healthy}, \text{sick first week}, \text{sick second week}, \text{sick third week}\}$$

The loss functions are now (fixed costs for additional heating increase to 18€):

$$L(\theta|\delta = \text{normal heating}) = \begin{cases} 0€ \text{ if } \theta = \text{healthy} \\ 20€ \text{ if } \theta = \text{sick first week} \\ 40€ \text{ if } \theta = \text{sick second week} \\ 60€ \text{ if } \theta = \text{sick first week} \end{cases}$$

$$L(\theta|\delta = \text{additional heating}) = \begin{cases} 18€ \text{ if } \theta = \text{healthy} \\ 38€ \text{ if } \theta = \text{sick first week} \\ 58€ \text{ if } \theta = \text{sick second week} \\ 78€ \text{ if } \theta = \text{sick first week} \end{cases}$$

The probabilities for the cases are:

state / decision	normal heating	additional heating
healthy	0.7^3=0.343	0.9^3=0.729
sick first week	0.3	0.1
sick second week	0.7*0.3=0,21	0.9*0.1=0.09
sick third week	0.7^2*0.3=0.147	0.9^2*0.1=0.081

This yields the risk:

$$\begin{aligned} R(\theta|\delta = \text{normal heating}) &= 0.3 \cdot 60€ + 0.21 \cdot 40€ + 0.147 \cdot 20€ = 29.34€ \\ R(\theta|\delta = \text{additional heating}) &= 0.729 \cdot 18€ + 0.1 \cdot 78€ + 0.09 \cdot 58€ + 0.081 \cdot 38€ \\ &= 29,22€ \end{aligned}$$

This time, because the cost of sickness increases faster than the cost of preventing sickness, additional heating is the slightly better option in this case.

Exercise 2

1.)

Firstly the decision space of estimating μ is $\mathbb{R}$. The lossfunction (if the loss is absolute) is given by:

$$L(\mu|\hat{\mu}) = |\hat{\mu} - \mu|$$

Therefore the loss for the 3 estimates are given by:

$$|1 - \mu|, |3 - \mu|, |5 - \mu|$$

Since the true value μ is not given, we can't compute the true values of the loss.

2.)

If we want to compute the Risk of the absolute loss meaning we are trying to solve:

$$R(\hat{\mu}|\mu) = \mathbb{E}(L(\hat{\mu}|\mu)) = \int_{\mathbb{R}} |\mu - \hat{\mu}| f_{\bar{X}_n}(\hat{\mu}) d\hat{\mu}$$

Using that $f_{\bar{X}_n}(x)$ is point symmetrical around μ, it holds that:

$$\begin{aligned}\int_{\mathbb{R}} |\mu - \hat{\mu}| f_{\bar{X}_n}(\hat{\mu}) d\hat{\mu} &\underset{hint1}{=} \int_{-\infty}^{\mu} (\mu - \hat{\mu}) f_{\bar{X}_n}(\hat{\mu}) d\hat{\mu} + \int_{\mu}^{\infty} (\hat{\mu} - \mu) f_{\bar{X}_n}(\hat{\mu}) d\hat{\mu} \\ &= 2\int_{\mu}^{\infty} (\hat{\mu} - \mu) f_{\bar{X}_n}(\hat{\mu}) d\hat{\mu} \\ &= 2\int_{\mu}^{\infty} (\hat{\mu} - \mu) \frac{\sqrt{n}}{\sqrt{2\pi}} \exp\left(-\frac{(\hat{\mu} - \mu)^2 n}{2}\right)\end{aligned}$$

Substituting $z = \hat{\mu} - \mu$ leads to:

$$\begin{aligned}2\int_{\mu}^{\infty} (\hat{\mu} - \mu) \frac{\sqrt{n}}{\sqrt{2\pi}} \exp\left(-\frac{(\hat{\mu} - \mu)^2 n}{2}\right) &\underset{\text{substitution}}{=} 2\int_{0}^{\infty} z \frac{\sqrt{n}}{\sqrt{2\pi}} \exp\left(-\frac{z^2 n}{2}\right) dz \\ &\underset{\text{symmetry around 0}}{=} \int_{-\infty}^{\infty} z \sqrt{\frac{n}{2\pi}} \exp\left(-\frac{z^2 n}{2}\right) dz\end{aligned}$$

But now it holds that:

$$\sqrt{\frac{n}{2\pi}} \exp\left(-\frac{z^2 n}{2}\right) = f_{\mathcal{N}}(z|0, 1/n)$$

Where $f_{\mathcal{N}}(z|0, 1/n)$ is the density of a normal distribution with mean 0 and variance $1/n$. $Z \sim \mathcal{N}(0, 1/n)$. But then:

$$0 = \mathbb{E}(Z) = \int_{-\infty}^{\infty} z f_{\mathcal{N}}(z|0, 1/n) dz = \int_{-\infty}^{\infty} z \sqrt{\frac{n}{2\pi}} \exp\left(-\frac{z^2 n}{2}\right) dz$$

So:

$$R(L(\hat{\mu}|\mu)) = \mathbb{E}(Z) = 0$$

Exercise 3

Mathematically the problem should be expressed as:

$$d_{\min} := \arg\min_{d \in \mathbb{R}} \mathbb{E}(L(\theta, d)) = \arg\min_{d \in \mathbb{R}} \int_{\mathbb{R}} w(\theta)(\theta - d)^2 f(\theta|x) d\theta$$

Where $f_\Theta(\theta)$ is the assumed pdf of θ. So the FOC should yield:

$$\frac{\partial}{\partial d} \int_{\mathbb{R}} w(\theta)(\theta - d)^2 f(\theta|x) d\theta \stackrel{!}{=} 0$$

Using the hint we can switch the derivative and integral and get:

$$\begin{aligned}
\int_{\mathbb{R}} \frac{\partial}{\partial d} w(\theta)(\theta - d)^2 f(\theta|x) d\theta &\stackrel{!}{=} 0 \\
\iff \quad -2 \int_{\mathbb{R}} w(\theta)(\theta - d) f(\theta|x) d\theta &= 0 \\
\iff \quad \int_{\mathbb{R}} w(\theta)(\theta - d) f(\theta|x) d\theta &= 0 \\
\iff \quad \int_{\mathbb{R}} w(\theta)\theta f(\theta|x) d\theta - d \int_{\mathbb{R}} w(\theta) f(\theta|x) d\theta &= 0 \\
\iff \quad \mathbb{E}(w(\theta)\theta|x) - d\mathbb{E}(w(\theta)|x) &= 0 \\
\iff \quad d &= \frac{\mathbb{E}(w(\theta)\theta|x)}{\mathbb{E}(w(\theta)|x)}
\end{aligned}$$

Meaning that the optimal decision $d_{\min}$ is the ratio of two expected values:

$$d_{min} := \frac{\mathbb{E}(w(\theta)\theta|x)}{\mathbb{E}(w(\theta)|x)}$$