Problem 3.1

1.

First state the general problem:

$$\hat{ heta}_{ML} = rg \max_{ heta \in \Theta} \mathcal{L}(x| heta) = rg \max_{ heta \in \Theta} \ln \mathcal{L}(x| heta)$$

Maximum likelihood function:

$$egin{aligned} \mathcal{L}(x| heta) &= \prod_{i=1}^n f(x_i| heta) \ &= \prod_{i=1}^n rac{1}{ heta} \mathrm{exp}\left(-rac{x_i}{ heta}
ight) \ &= heta^{-n} \mathrm{exp}\left(-rac{1}{ heta} \sum_{i=1}^n x_i
ight) \end{aligned}$$

Now log likelihood:

$$\ln \mathcal{L}(x| heta) = -n \ln heta - rac{1}{ heta} \sum_{i=1}^n x_i$$

FOC:

$$egin{aligned} rac{\partial \ln \mathcal{L}(x| heta)}{\partial heta} &= -rac{n}{ heta} + rac{1}{ heta^2} \sum_{i=1}^n x_i = 0 \ rac{1}{ heta} \sum_{i=1}^n x_i &= n \ heta_{ML} &= ar{X}_n \end{aligned}$$

SOC:

$$egin{align} \ln \mathcal{L}''(x| heta) &= rac{n}{ heta^2} - rac{2}{ heta^3} \sum_{i=1}^n x_i \ &= rac{n}{ heta^2} - rac{2n}{ heta^3} ar{X}_n \end{split}$$

Means, if we check at $x = \bar{X}_n$:

$$egin{align} \ln \mathcal{L}''(ar{X}_n| heta) &= rac{n}{(ar{X}_n)^2} - rac{2n}{(ar{X}_n)^3}ar{X}_n \ &= rac{n-2n}{ar{X}_n^2} < 0 \end{split}$$

2.

Unbiasedness $\mathbb{E}(ar{X}_n) = heta$:

$$egin{aligned} \mathbb{E}(ar{X}_n) &= \mathbb{E}\left(rac{1}{n}\sum_{i=1}^n X_i
ight) \ &= rac{1}{n}\sum_{i=1}^n \mathbb{E}(X_i) \ &= rac{1}{n}\sum_{i=1}^n heta \ &= rac{n}{n} heta = heta \end{aligned}$$

Variance:

$$egin{split} Var(ar{X}_n) &= rac{1}{n^2} Var\left(\sum_{i=1}^n X_i
ight) \ &= rac{1}{n^2} \sum_{i=1}^n Var(X_i) \quad ext{iid} \ &= rac{n}{n^2} heta^2 = rac{ heta^2}{n} \end{split}$$

3.

Bias variance decomposition:

$$egin{aligned} R(\delta, heta) &= \mathbb{E}((\delta(X) - heta)^2) \ &= Bias(\delta)^2 + Var(\delta) \end{aligned}$$

But as we look only at unbiased estimators our risk becomes:

$$R(\delta, \theta) = Var(\delta)$$

And because $Var(\bar{X}_n)=\frac{\theta^2}{n}$ and therefore has the lowest variance, it also has the lowest risk and is therefore admissible.

4.

First compute the bias:

$$egin{aligned} \mathbb{E}(\hat{ heta}_2) &= \mathbb{E}\left(rac{1}{n+1}\sum_{i=1}^n X_i
ight) \ &= rac{1}{n+1}\sum_{i=1}^n \mathbb{E}(X_i) \ &= rac{n}{n+1} heta \ \implies Bias(\hat{ heta}_2) &= rac{n}{n+1} heta - heta \ &= -rac{1}{n+1} heta \end{aligned}$$

Then compute the variance

$$egin{align} Var(\hat{ heta}_2) &= rac{1}{(n+1)^2} \sum_{i=1}^n Var(X_i) \ &= rac{n}{(n+1)^2} heta^2 \ \end{aligned}$$

And finally combine in the Bias variance decomposition:

$$egin{aligned} R(\hat{ heta}_2, heta) &= Bias(\hat{ heta}_2)^2 + Var(\hat{ heta}_2) \ &= rac{1}{(n+1)^2} heta^2 + rac{n}{(n+1)^2} heta^2 \ &= rac{(n+1)}{(n+1)^2} heta^2 \ &= rac{1}{n+1} heta^2 < rac{1}{n} heta^2 = R(L(\hat{ heta}_{ML}, heta)) \end{aligned}$$

So even-though the max. likelihood est. is an admissible unbiased estimator, the estimator is not overall admissible.

Problem 3.2

1.

$$egin{aligned} \Theta &= (0,1) \ \mathscr{X} &= \{0,\ldots,k\}^n \ D &= \Theta \end{aligned} \ L &= L(heta,d) = \min \left\{ 2, \left(rac{d- heta}{ heta}
ight)^2
ight\}$$

Let $\delta_M \in D$ a decision rule, then δ_M is minimax iff:

$$\sup_{\theta \in \Theta} R(\delta, \theta) = \inf_{\delta \in D} \sup_{\theta \in \Theta} R(\delta, \theta)$$

Where R is the expected loss under parameter θ .

3.

$$L(heta,\delta_0) = \min \left\{ 2, \left(-rac{ heta}{ heta}
ight)^2
ight\} = 1 \quad orall heta \in \Theta$$

If we were to choose another randomized rule, this rule would need to be positive $\delta(X)=c>0.$

In that case it would hold, that:

$$g(x, heta) = \left(rac{c- heta}{ heta}
ight)^2 = \left(rac{c}{ heta} - rac{ heta}{ heta}
ight)^2 = \left(rac{c}{ heta} - 1
ight)^2$$

In that case we can choose $\theta \to 0$ and get $g(x,\theta) \to \infty$ and therefore a loss $R(\theta,\delta) = L(\theta,\delta) = 2$.

So there is no other rule, with a smaller "worst case" scenario.

Problem 3.3

a)

$$egin{aligned} \Theta &= H_0 \cup H_1 = \mathbb{R}_+ \ \mathscr{X} &= \mathbb{R}^n \ D &= \{\delta : \mathscr{X}
ightarrow \{0,1\}\} \ L &= 0\text{-}1 ext{ loss} \end{aligned}$$

Because σ is given, we can assume, that the parameter space is just the space of μ .

b)

Do we need to show this?

Assuming iid we can say (because any linear combination of normally distributed random variables is also normal):

$$ar{X}_n = rac{1}{n} \sum_{i=1}^n X_i \sim \mathcal{N}\left(\mu, rac{\sigma^2}{n}
ight)$$

Likewise we can say:

$$ar{X}_n - \mu \sim \mathcal{N}\left(0, rac{\sigma^2}{n}
ight)$$

And finally:

$$Z_n = rac{(ar{X}_n - \mu)\sqrt{n}}{\sigma} \sim \mathcal{N}(0,1)$$

Because:

$$egin{aligned} Var(Z_n) &= Var\left(rac{\sqrt{n}}{\sigma}(ar{X}_n - \mu)
ight) \ &= rac{n}{\sigma^2}Var(ar{X}_n - \mu) \ &= rac{n}{\sigma^2}Var(ar{X}_n) \ &= rac{n}{\sigma^2} \cdot rac{\sigma^2}{n} = 1 \end{aligned}$$

since we know, that $\sigma^2=4$ we can now make the statement:

$$P\left(Z_n \geq q_{0.975}
ight) = 2.5\% \ \iff P\left(Z_n \geq 1.96
ight) = 2.5\% \ \iff P\left(rac{(ar{X}_n - \mu)\sqrt{n}}{2} \geq 1.96
ight) = 2.5\% \ \iff P_{H_0}\left(rac{ar{X}_n\sqrt{n}}{2} \geq 1.96
ight) = 2.5\%$$

Because this exactly equivalent to the critical region C_n of our test we have now computed the probability for a type 1 error, i.e. the significance level. This means, that the significance level of the test is equal to 2.5%.

c)

We are looking the solution to the following term:

$$P_{H_1}(t(ec{X}) \geq 1.96) = 1 - P_{H_1}\left(rac{ar{X}_n\sqrt{n}}{2} < 1.96
ight)$$

Meaning the test statistic realizes a value outside of the critical region under the alternative. Under the alternative the term has the distribution:

$$t(ec{X}) \sim \mathcal{N}(\mu, 1)$$

meaning that we can describe the probability of a type 2 error like so:

$$1-\Phi\left(t(ec{X})-rac{\mu\sqrt{n}}{2}
ight)$$

So the power is dependent on the size of μ . The higher μ the higher the power.

d)

$$egin{aligned} \lambda(ec{X}) &= rac{\mathcal{L}(ec{X},0)}{\sup_{\mu \in (0,\infty)} \mathcal{L}(ec{X},\mu)} \ &= rac{\left[rac{1}{\sqrt{2\pi}\sigma}
ight]^n}{\left[rac{1}{\sqrt{2\pi}\sigma}
ight]^n} \mathrm{exp}\left(rac{1}{2\sigma^2}\sum_{i=1}^n x_i^2 - rac{1}{2\sigma^2}\sum_{i=1}^n (x_i - ar{x}_n)^2
ight) \ &= \mathrm{exp}\left(2\sum_{i=1}^n x_iar{x}_n - n
ight) \end{aligned}$$

e)

Question: is the likelihood ratio montonic?

f)

We can express the risk depending on weither H_0 is true, or the alternative:

$$R(t,\mu) = egin{cases} 2.5\% & ext{if } \mu = 0 \ 1 - \Phi\left(t(ec{X}) - rac{\mu\sqrt{n}}{2}
ight) & ext{else} \end{cases}$$

and yes the function is piecewise monotonic, because it jumps to 97,5% at $\mu \to 0$ and then decreases. For $\mu=0$ it is a single value.