

EXAM 2025

Problem 1

a. - We look at p -values from Wald tests, and based on these, the following covariates are very important: season, hour, daynight, weather, temperature and humidity.

Less important: holiday and windspeed

Not important: workingday

- The effect of temp, which is 4,33, is the expected additional number of rented bikes when the temperature increases by 1°C , all the other covariates remaining the same.
- For hr and daynight, it is not always possible to increase one of them without affecting the other.

b. - $AIC = -2 \log \text{lik}(\hat{\theta}) + 2 |\hat{\theta}|$

$BIC = -2 \log \text{lik}(\hat{\theta}) + \log(N) |\hat{\theta}|$

$|\hat{\theta}|$ number of free parameters

BIC is larger because it puts a higher penalty on $|\hat{\theta}|$.

- The removed covariates are the ones that were the least important in Model 1.
- If prediction performance is the most important, you would tend to put most weight on the AIC and choose Model 2, whereas if you prefer a simpler model, for instance for interpretation, you would tend to put more weight on the BIC and choose Model 3.

d. - The number of rented bikes increases during the day up to approx. 8 p.m., and then decreases.

The wind speed does not affect the number of rented bikes if it is below approx. 25 km/h, but for higher values it has a negative effect.

- According to the table on nonparametric ANOVA, all effects of the numerical covariates are significantly non-linear (low p -values) and the AIC and BIC are much lower for the additive model than for the linear ones, indicating that the model is improved by including non-linear effects.

e. - The important variables are ones used in the splits of the tree: hr, kmh, season and workingday.