

# Assignment 3

## STAT40850 Bayesian Analysis (online)

2024/2025

AUTHOR

Alberto Caimo

## Instructions

---

- Hand-in date: **Wednesday 2nd April 2025 at 11:59pm.**
- You should submit it to the Assignment 3 object in Brightspace.
- You should submit a pdf file and corresponding source file (e.g., rmd or qmd file) containing your answers to the questions.
- You may submit it multiple times before the deadline, but only the last version will be marked.

## Dataset

---

These data are the results of a chemical analysis of 178 wines grown in the same region in Italy. The analysis determined the quantities of 4 constituents found in all the types of wine.

To load the dataset:

```
load(url("https://acaimo.github.io/teaching/data/italian_wines.RData"))
head(italian_wines)
```

	alcohol	magnesium	color_intensity	proline
1	14.23	8.892011	5.64	3.381946
2	13.20	7.001583	4.38	3.334313
3	13.16	7.071599	5.68	3.763010
4	14.37	7.911789	7.80	4.699793
5	13.24	8.261868	4.32	2.334019
6	14.20	7.841773	6.75	4.604527

## Questions

---

1. Use Stan to perform Bayesian inference by implementing **three competing linear (Gaussian)** regression models to predict the wine's **alcohol** level based on the available covariate information. Justify the choice of prior distributions for each model, explaining the underlying rationale. After fitting the models, **summarise the posterior parameter distributions and interpret the results for each case.**

2. Employ two different posterior predictive checks of your choice to examine the model fit for each model. Provide a concise interpretation of the results obtained.

[20]

3. Compare the three models using WAIC and LOO-cross-validation. Provide a concise interpretation.

[20]

4. Propose an alternative **non-Gaussian** model (that is appropriate for predicting the target variable, **alcohol**) based on the same covariate included in the best model selected previously. Justify your choice of the model likelihood and prior distributions. Compare the performance of your model to the best model identified in Question 3. Offer a clear and concise interpretation of posterior results and model comparison outcome.

[30]