

# Practice Problems on Bayesian Logistic Regression

## Q1

You are building a Bayesian logistic regression model to predict whether a young adult will vote. You include two features: **age** and **education level (in years)**. You place a prior belief that young age reduces the chance of voting, and encode this with a prior:

- $w_{\text{age}} \sim \mathcal{N}(-0.8, 0.3^2)$
- Initialize  $w_{\text{edu}} = 0.5$

The model uses the logistic function:

$$P(y = 1 \mid x) = \frac{1}{1 + \exp(-w^T x)}$$

- a) A new user has age = 22 and education = 16. Assume the MAP estimate for  $w_{\text{age}} = -0.6$  and keep  $w_{\text{edu}} = 0.5$ . Calculate the probability this user will vote using the logistic function.
- b) If the model had no prior on  $w_{\text{age}}$ , the MAP estimate would have been  $w_{\text{age}} = -0.2$ . Would this change increase or decrease the predicted probability? Why?

## Q2

You're designing a wellness app that predicts if a user is "at risk" (1) or "not at risk" (0) for high stress based on hours of sleep and daily screen time. The app developers believe that less sleep strongly increases risk.

- a) If you set a prior on the coefficient for "hours of sleep" as  $w_{\text{sleep}} \sim \mathcal{N}(-1.0, 0.5^2)$ , what does this mean in plain English? Explain the role of this prior in influencing model predictions.
- b) Write 3 lines of pseudocode to:

- Apply sigmoid to weighted input  $w^T x$
- Multiply by label  $y$  to get likelihood
- Multiply by prior (in log-space)

### Q3

You have built a Bayesian logistic regression model to classify emails as spam (1) or not spam (0). You want to flag emails for review if the model isn't confident.

- a) An email gets a predictive probability of 0.5 for being spam. What does this tell you about the model's confidence? Should it be flagged?
- b) If your model has no prior information, it behaves like standard logistic regression. How does adding a prior on weights help when you have only 100 labeled emails? Answer in 2 sentences.