

Mathematical details of study *Too Tired to Know*

Alvin Gavel

1 Introduction

In the article *Too Tired to Know* we implement the cumulative model described in Bürkner and Vuorre (2019). The reader is primarily recommended to that article for the details, or – once it is written – the section in our article that explains the analysis. This is mostly written for myself to force myself to understand how the model works, since the root of at least half of all evil is the use of statistical recipes without first understanding them. The model is implemented in Python rather than R for the same reason, so that instead of copy-and-pasting their R code I have to think about how to get the equivalent Python code¹.

2 Basic idea

The cumulative model starts out from the assumption that the ordinal variable that we actually measure, Y , is an increasing function of some underlying unobservable variable \tilde{Y} . Specifically it assumes that there are K thresholds τ_k such that if $\tau_{k-1} < \tilde{Y} < \tau_k$ then we will observe $Y = k$. The underlying \tilde{Y} in turn follows some probability density function f such that the probability $P(Y = k)$ is given by

$$P(Y = k) = \int_{\tau_{k-1}}^{\tau_k} f(\tilde{Y}) d\tilde{Y} \quad (1)$$

If we call the corresponding cumulative density function F , this is

$$= F(\tau_k) - F(\tau_{k-1}) \quad (2)$$

3 Linear model of prediction term

So far the model is very general, but to get anywhere we need to make some assumptions about what the function f actually looks like. We split \tilde{Y} into a predictor term η and an error term ε , so that

$$\tilde{Y} = \eta + \varepsilon \quad (3)$$

¹Plus, I just don't like R.

We now assume that η is a linear function, so that

$$\eta = \sum_{n=0}^K b_n x_n \quad (4)$$

Inserting (4) into (3) gives us

$$\tilde{Y} = \sum_{n=0}^K b_n x_n + \varepsilon \quad (5)$$

Since $d\tilde{Y}/d\varepsilon = 1$ we can rewrite (1) as

$$P(Y = k) = \int_{\tau_{k-1}}^{\tau_k} f\left(\sum_{n=0}^K b_n x_n + \varepsilon\right) d\varepsilon \quad (6)$$

Shifting the integration boundaries by the predictor term gives us

$$= F\left(\tau_k - \sum_{n=0}^K b_n x_n\right) - F\left(\tau_{k-1} - \sum_{n=0}^K b_n x_n\right) \quad (7)$$

4 Normal model of error term

We now need some kind of model of what that error term looks like. To make life simpler², we assume that it is normally distributed. If we denote the CDF of the normal distribution with Φ , then (7) takes the form:

$$P(Y = k) = \Phi\left(\tau_k - \sum_{n=0}^K b_n x_n\right) - \Phi\left(\tau_{k-1} - \sum_{n=0}^K b_n x_n\right) \quad (8)$$

This means we have a regression problem where we need to estimate the thresholds τ_k and τ_{k-1} and the regression coefficients b_n .

²It is common to justify assumptions of normality with reference to the central limit theorem. I would argue that this often over-interprets the central limit theorem as being much broader than it actually is, and that the real reason why normal distributions get used so often is that they are just so very convenient.