Université des Sciences et Techniques de Nantes

Mars 2023

Etude des germinations par une régression logistique à effet aléatoire

MINETTE Cloée, BAKHOUCHE Sarah,
SIMON Antoine, DAVID Romain

Table des matières

1 Introduction

Le jeu de données SEEDS a été étudié par Crowder en 1978. Le but est de modéliser la germination de deux types de graines, *O. aegyptiaco 75* et *O. aegyptiaco 73* pour deux variétés différentes, concombre et haricot. Nous regardons la germination de ces graines pour chacune des 21 plaques. tableau des données + représentation graphique pourquoi modèle logistique ?

2 Modèle

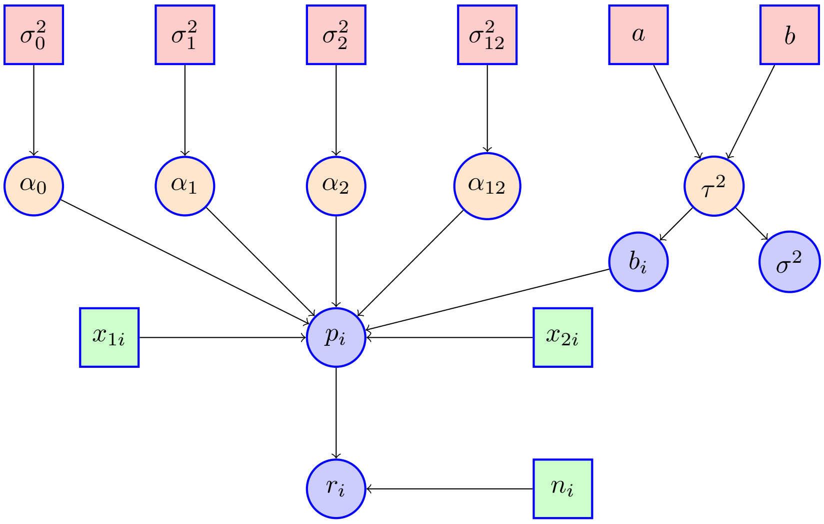
2.1 Choix du modèle

Notre jeu de données compte le nombre de graines germées sur chaque plaque. Nous sommes donc soumis à un espace discret. Ces données de dénombrement sont souvent influencée par une variation des données plus grande que prévue, c'est la surdispersion. Pour cette étude, il est suggérable d'utiliser un modèle bayésien Binomial hiérarchique à effet aléatoire afin de mieux capturer la surdispersion. C'est pourquoi on décide d'appliquer ce modèle au jeu de données SEEDS. Il est exprimé comme suit, avec r_i le nombre de graines germées sur la i-ème plaque, n_i le nombre total de graines sur la i-ème plaque, p_i la probabilité de germination sur la i-ème plaque, où x_{1i} et x_{2i} sont le type de semence et l'extrait de racine de la plaque i. Nous ajoutons à celà le terme d'intéraction $\alpha_{12}x_{1i}x_{2i}$.
Pour $i \in \{1; 21\}$, le modèle est tel que :

$r_i \sim Binomiale(p_i, n_i)$
$logit(p_i) = \alpha_0 + \alpha_1 x_{1i} + \alpha_2 x_{2i} + \alpha_{12} x_{1i} x_{2i} + b_i$
$b_i \sim Normal(0, \tau)$
explication des paramètres et des lois

2.2 Construction du modèle

On peut représenter notre modèle par un graphe orienté acyclique (Directed Acyclic Graphe (DAG)) :



$$\Pi(y_{ij}|...) \propto \Pi(y_{ij}|\mu_{ij}) \propto \sum_{k=0}^{k=\infty} \frac{\mu_{ij}^k e^{\mu_{ij}}}{k!}$$

$$\Pi(\mu_{ij}|...) \propto \Pi(\mu_{ij}|\alpha,\beta,\gamma,\lambda_{ij},x_{ij})\Pi(y_{ij}|\mu_{ij}) \propto e^{\alpha+\beta log(x_{ij}+10)+\gamma x_{ij}+\lambda_{ij}} \sum_{k=0}^{k=\infty} \frac{\mu_{ij}^k e^{\mu_{ij}}}{k!}$$

$$\Pi(x_{ij}|...) \propto \Pi(x_{ij})\Pi(\mu_{ij}|x_{ij},\lambda_{ij},\beta,\gamma,\alpha) \propto \Pi(x_{ij})e^{\alpha+\beta log(x_{ij}+10)+\gamma x_{ij}+\lambda_{ij}}$$

$$\Pi(\lambda_{ij}|...) \propto \Pi(\lambda_{ij}|\tau)\Pi(\mu_{ij}|\lambda_{ij},\beta,\gamma,\alpha,x_{ij}) \propto \frac{1}{\sqrt{2\pi\tau}}e^{-\frac{\lambda_{ij}^2}{2\tau^2}}$$

$$\begin{aligned}\Pi(\alpha|...) &\propto \Pi(\alpha)\prod_{ij}\Pi(\mu_{ij}|\alpha,\beta,\gamma,\lambda_{ij},x_{ij}) \propto \frac{1}{\sqrt{2\pi\sigma_\alpha^2}}e^{-\frac{\alpha^2}{2\sigma_\alpha^2}}\prod_{ij}e^{\alpha+\beta log(x_{ij}+10)+\gamma x_{ij}+\lambda_{ij}} \\ &\propto \frac{1}{\sigma_\alpha}e^{-\frac{\alpha^2}{2\sigma_\alpha^2}+n\alpha+\sum_{ij}\beta log(x_{ij}+10)+\sum_{ij}\gamma x_{ij}+\sum_{ij}\lambda_{ij}}\end{aligned}$$

écriture des lois a priori a posteriori conditionnelle

3 Analyse et interprétation des résultats