Generalized linear models

Jean-Michel Marin

University of Montpellier Faculty of Sciences

HAX912X - 2025/2026

- Why generalized linear models
- Scalar exponential families
- Exponential families with a nuisance parameter
- Definition of generalized linear models
- 6 Classical examples

In many applications, the response does not vary in all $\mathbb R$ but in $\mathbb R^+,$ in $\mathbb N,$ in $\{0,1\}...$

The Gaussian model is not suited to this situation

 $y = (y_1, \dots, y_n)$ the vector of responses X the matrix of explanatory variables

The distribution of y_i , $(\mathbb{P}_{\theta_i})_{\theta_i \in \mathbb{R}}$ must be specified $\mathscr{P}(\theta_i)$, $\mathscr{E}(\theta_i)$, $\mathscr{B}(\theta_i)$, $\mathscr{N}(\theta_i, 1)$, ...

The link between θ_i and X must also be specified

We assume that $\gamma(\theta_i) = x_i \beta$ $\gamma(\cdot)$ is called the link function

A GLM is fully specified by

- a probability family
- a link function

Gaussian linear model

$$\begin{split} \mathbb{P}_{\theta} &= \mathscr{N}(\theta, \sigma^2) \\ \gamma(x_i \beta) &= x_i \beta \end{split}$$



5/19

Jean-Michel Marin (FdS) GLM HAX912X

Examples

- Gaussian linear model
- Logistic regression model
- Poisson regression model

Let v(dx) be a reference measure on \mathbb{R} ,

$$b(\theta) = \log \left(\int \exp(\theta y) \nu(dy) \right)$$

and

$$D_{\nu} = \{\theta|b(\theta) < \infty\} \subseteq \mathbb{R}$$

Definition

A family of probability distribution \mathbb{P}_{θ} is said to belong to the scalar exponential family if

• for each element of the family there exist a $\theta \in D_{\nu}$ such that the probability distribution can be written in the form

$$\mathbb{P}_{\theta}(dx) = \exp(\theta x - b(\theta))\nu(dx)$$

b to any value of θ corresponds one and only one element of the family

 θ is called the natural parameter of the exponential family The exponential family is said to be regular if D_{ν} is open

8/19

If θ is an interior point of D_{ν} then

$$b'(\theta) = \mathbb{E}_{\theta}(y)$$

$$b''(\theta) = \mathbb{V}_{\theta}(y)$$

The function $b(\theta)$ is strictly convex

The strictly convex nature of $\mathfrak{b}(\theta)$ means that $\mathfrak{b}'(\theta)$ is bijective

We can also consider $\mu = \mathbb{E}_{\theta}(y)$ as a parameter

Examples

- Poisson distribution with parameter $\lambda > 0$
- ▶ Binomial distribution with parameters (m, p) where m is fixed and $p \in]0, 1[$
- \blacktriangleright Gaussian distribution with parameters (μ,σ^2) where σ^2 is known and $\mu\in\mathbb{R}$

Maximum likelihood estimation of θ

Let y_1, \ldots, y_n be an n-sample from \mathbb{P}_{θ^*}

If \mathbb{P}_{θ} belongs to the scalar exponential family with θ as the natural parameter, then $\hat{\theta}_n$ the MLE of θ^* is such that

$$\frac{1}{n}\sum_{i=1}^n y_i = b'(\hat{\theta}_n)$$

$$D_{\nu,\varphi} = \left\{\theta \left| \int \text{exp}\left[\frac{x\theta - b(\theta)}{\varphi} + c(x,\varphi)\right] \nu(dx) < \infty \right.\right\}$$

Definition

A family of probability distribution $\mathbb{P}_{(\theta,\varphi)}$ is said to belong to the exponential family with nuisance parameter φ if

for each element of the family there exist a $\theta \in D_{\nu,\varphi}$ and a $\varphi \in \mathbb{R}^+$ such that the probability distribution can be written in the form

$$\mathbb{P}_{\theta, \varphi}(\mathrm{d} x) = \exp\left\{\frac{x\theta - b(\theta)}{\varphi} + c(x, \varphi)\right\} \nu(\mathrm{d} x)$$

▶ to any pair of $\theta \in D_{\nu,\varphi}$ and $\varphi \in \mathbb{R}^+$ corresponds one and only one element of the family

Jean-Michel Marin (FdS) GLM HAX912X 13/19

We have

$$b'(\theta) = \mathbb{E}_{\theta}(y)$$
$$b''(\theta) = \frac{\mathbb{V}_{\theta}(y)}{\Phi}$$

Examples

- ▶ Gaussian distribution with parameters (μ, σ^2) where $\mu \in \mathbb{R}$ and $\sigma^2 \in \mathbb{R}^+$
- ▶ Gamma distribution with parameters $\alpha > 0$ and $\beta > 0$

Jean-Michel Marin (FdS) GLM

Maximum likelihood estimation of θ

Let y_1, \dots, y_n be an n-sample from $f(y; \theta^*, \varphi^*)\nu(dx)$

For any $\varphi^*,\,\hat{\theta}_n$ the MLE of θ^* is such that

$$\frac{1}{n}\sum_{i=1}^{n}y_{i}=b'(\hat{\theta}_{n})$$

Definition of generalized linear models

Consider the n-sample $(x_i, y_i)_{i=1,\dots,n}$ from (x, y) where x is the vector of explanatory variables and y the corresponding response

Definition

Choosing a generalized linear model corresponds to choosing a conditional probability distribution for y|x. For the class of generalized linear model this conditional distribution is such that

be the distribution of y|x belongs to an exponential family with a nuisance parameter

$$\gamma(\mathbb{E}(\mathbf{y}|\mathbf{x})) = \mathbf{x}\boldsymbol{\beta}$$

 $\gamma(\cdot)$ is called the link function



17/19

Jean-Michel Marin (FdS) GLM HAX912X

Definition of generalized linear models

- Choosing the exponential family determined in most cases by the values taken by y; if several choices are possible, the plots of the residuals can be used to decide which family is the most appropriate
- 2) Choice of link function: we can use the canonical link: $\gamma(\cdot) = b'(\cdot)$ in this case we have $\theta = x\beta$ that is a natural and advantageous choice, many formulas are simplified

18/19

Classical examples

Logistic regression

$$\begin{split} \mathbb{P}_{\theta} &= \mathscr{B}(\theta) \\ \gamma(\mathfrak{u}) &= \log(\mathfrak{u}/(1-\mathfrak{u})) \\ \mathbb{E}(\mathfrak{y}|\mathfrak{x}) &= \exp(\mathfrak{x}\beta)/(1+\exp(\mathfrak{x}\beta)) \end{split}$$

Poisson regression
$$\mathbb{P}_{\theta} = \mathscr{P}(\theta)$$
 $\gamma(\mathfrak{u}) = log(\mathfrak{u})$



