Generalised Additive Models for Location Scale and Shape

Mikis Stasinopoulos[1] Bob Rigby[1]

[1]STORM, London Metropolitan University

GAMLSS: Short course Cordoba, Argentina, Oct 2012

gamlss

Summary of the GAMLSS course

Day 1:

- Introduction to GAMLSS
- introduction to the GAMLSS software

Day 2:

- Continuous distributions in GAMLSS
- Discrete, mixed and finite mixture distributions in GAMLSS

Day 3:

- Smoothing and random effects within GAMLSS
- Centile estimation and further work

gamlss

gamlss

The Dutch boys data

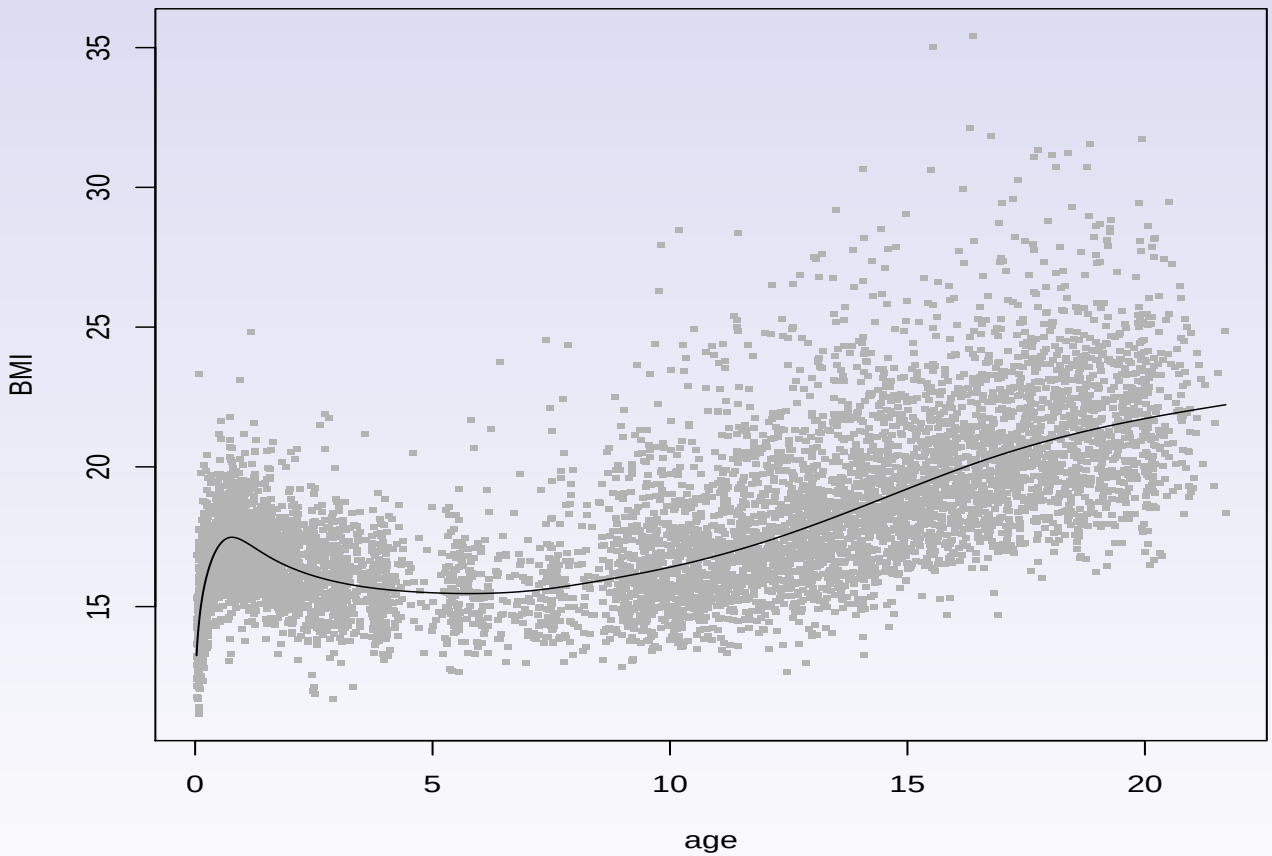
`BMI` : the BMI of 7294 boys

`age` : the age in years

Source: Buuren and Fredriks (2001)

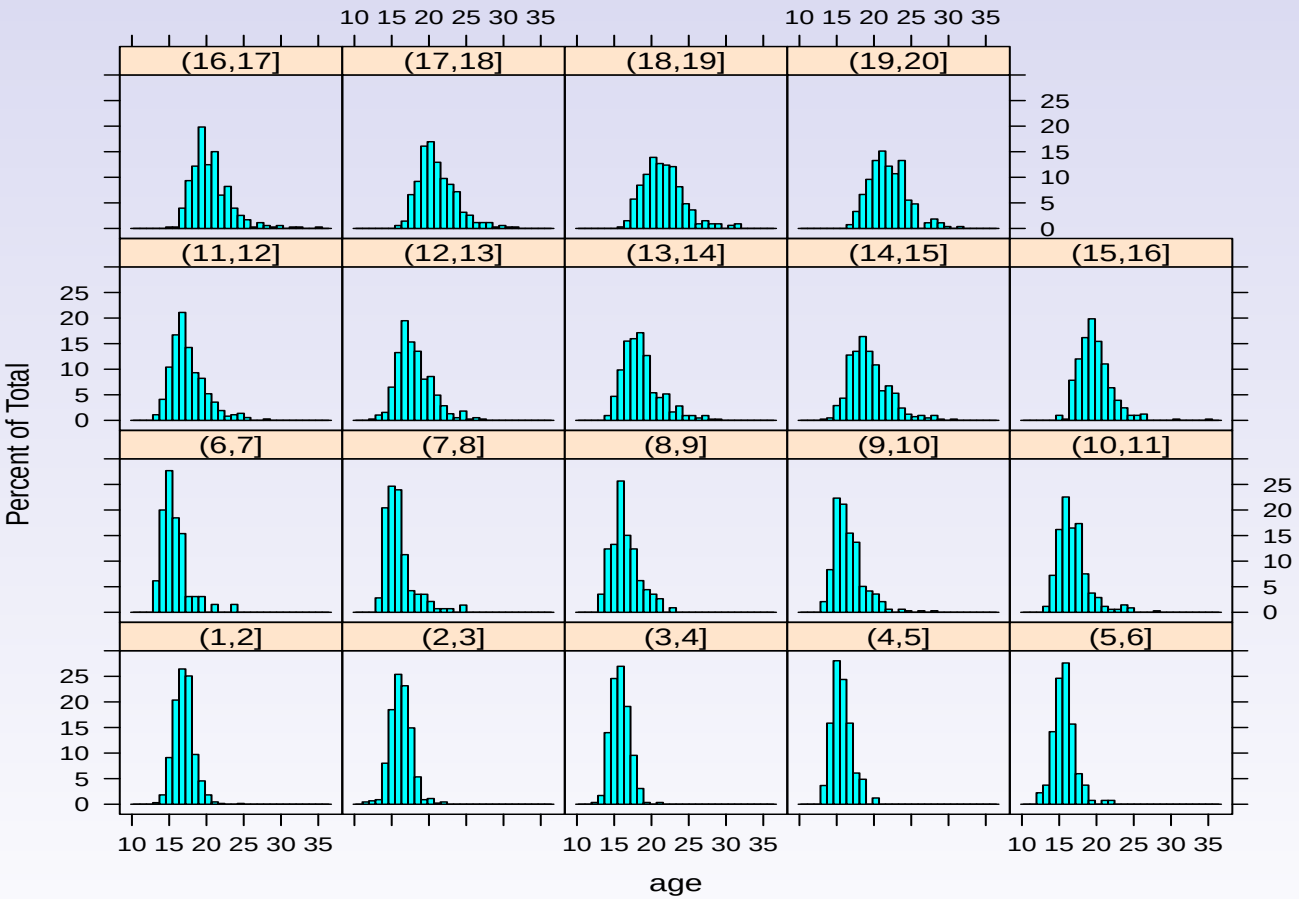
gamlss

The Dutch boys data: statistical challenges



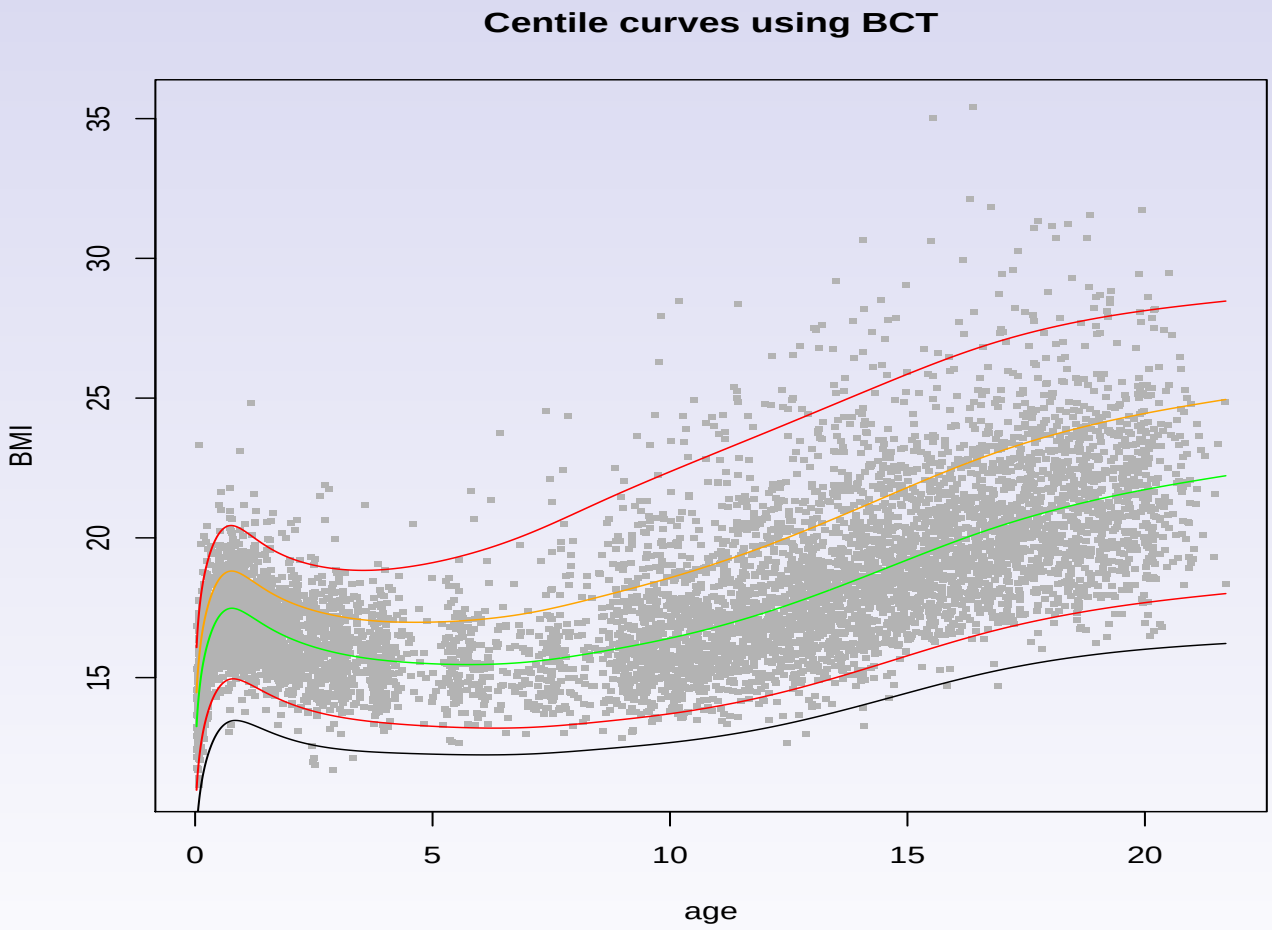
gamlss

The Dutch boys data: Histograms by age



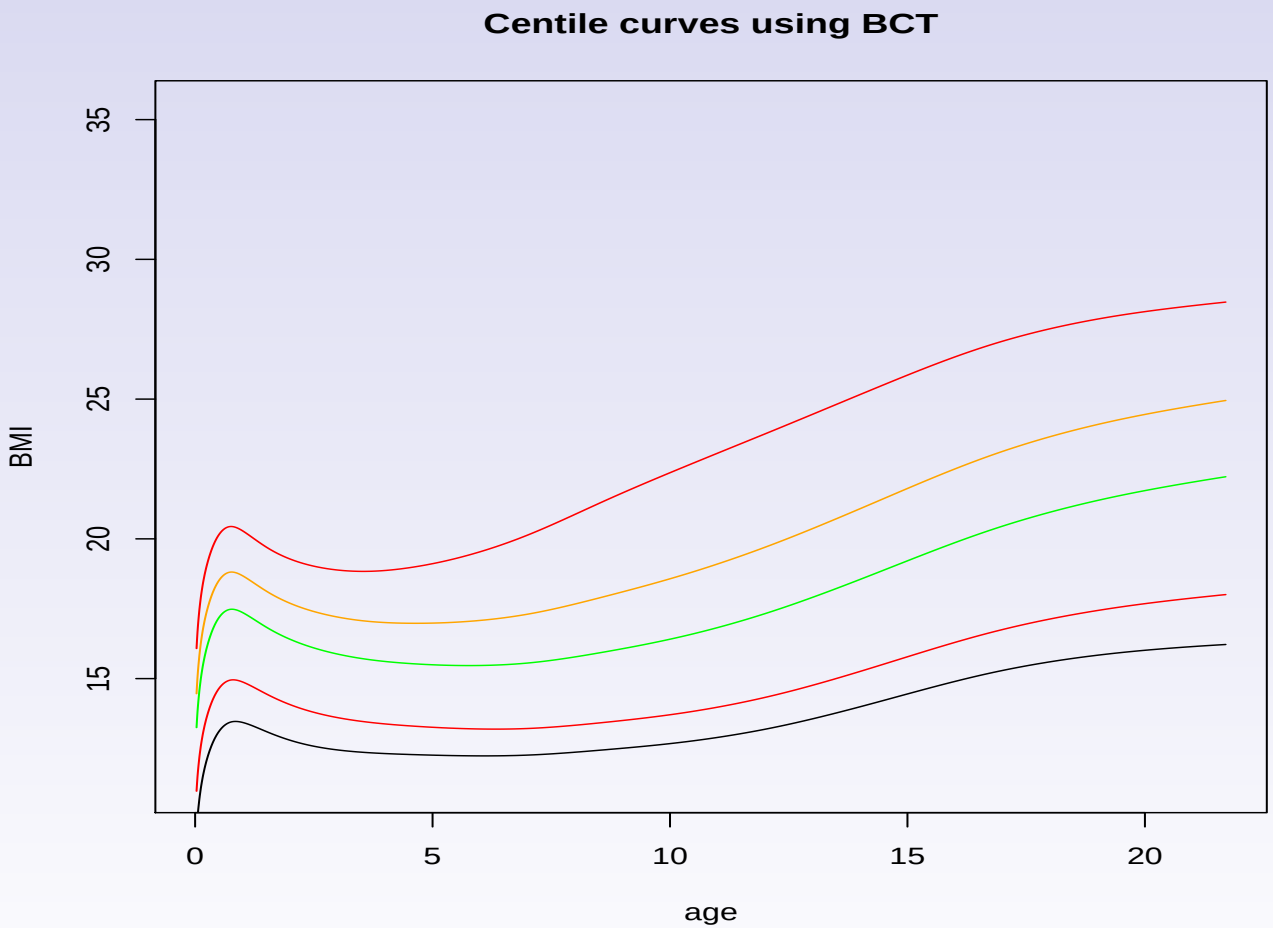
gamlss

The Dutch boys data: centile estimation



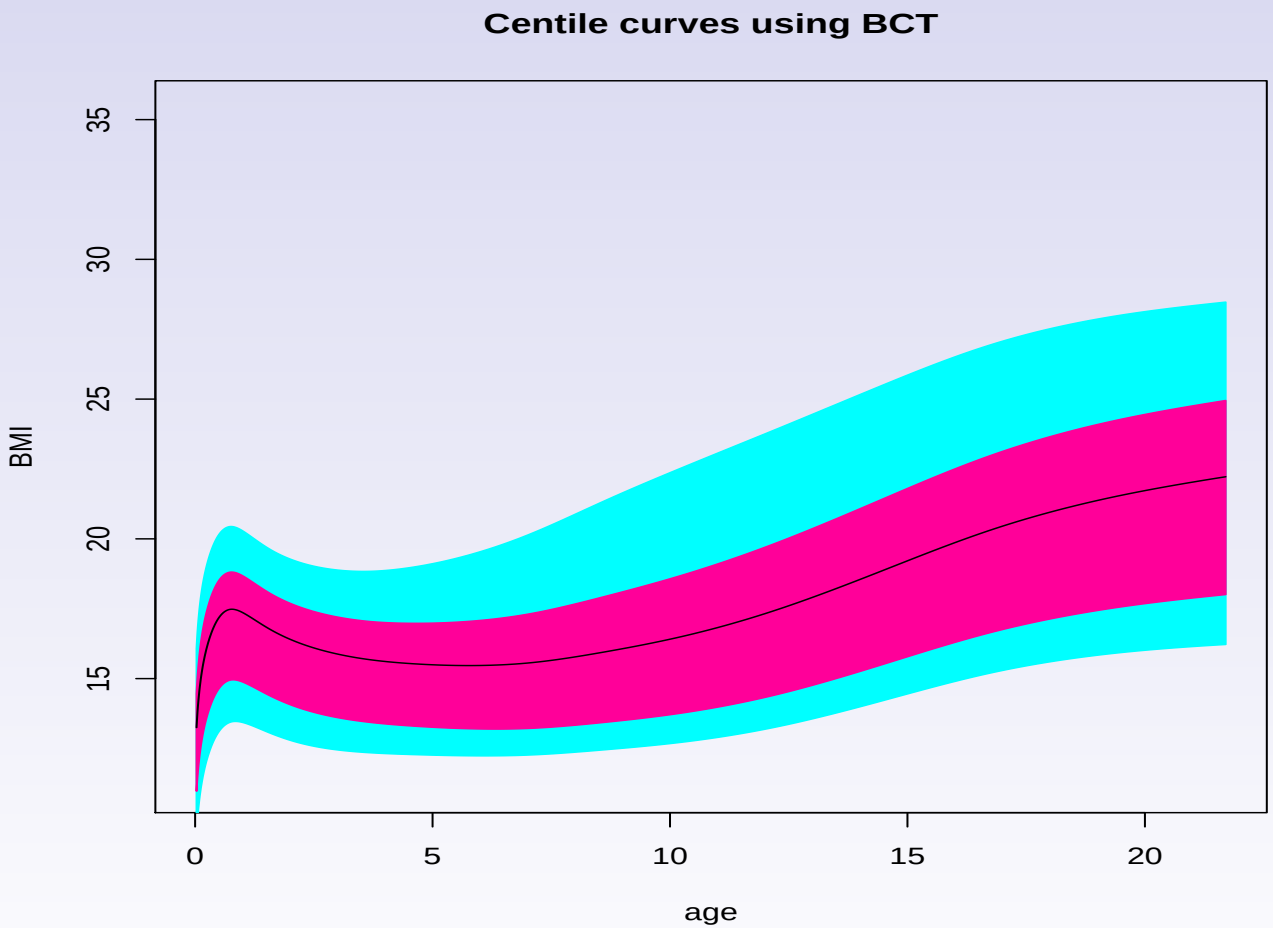
gamlss

The Dutch boys data: centiles



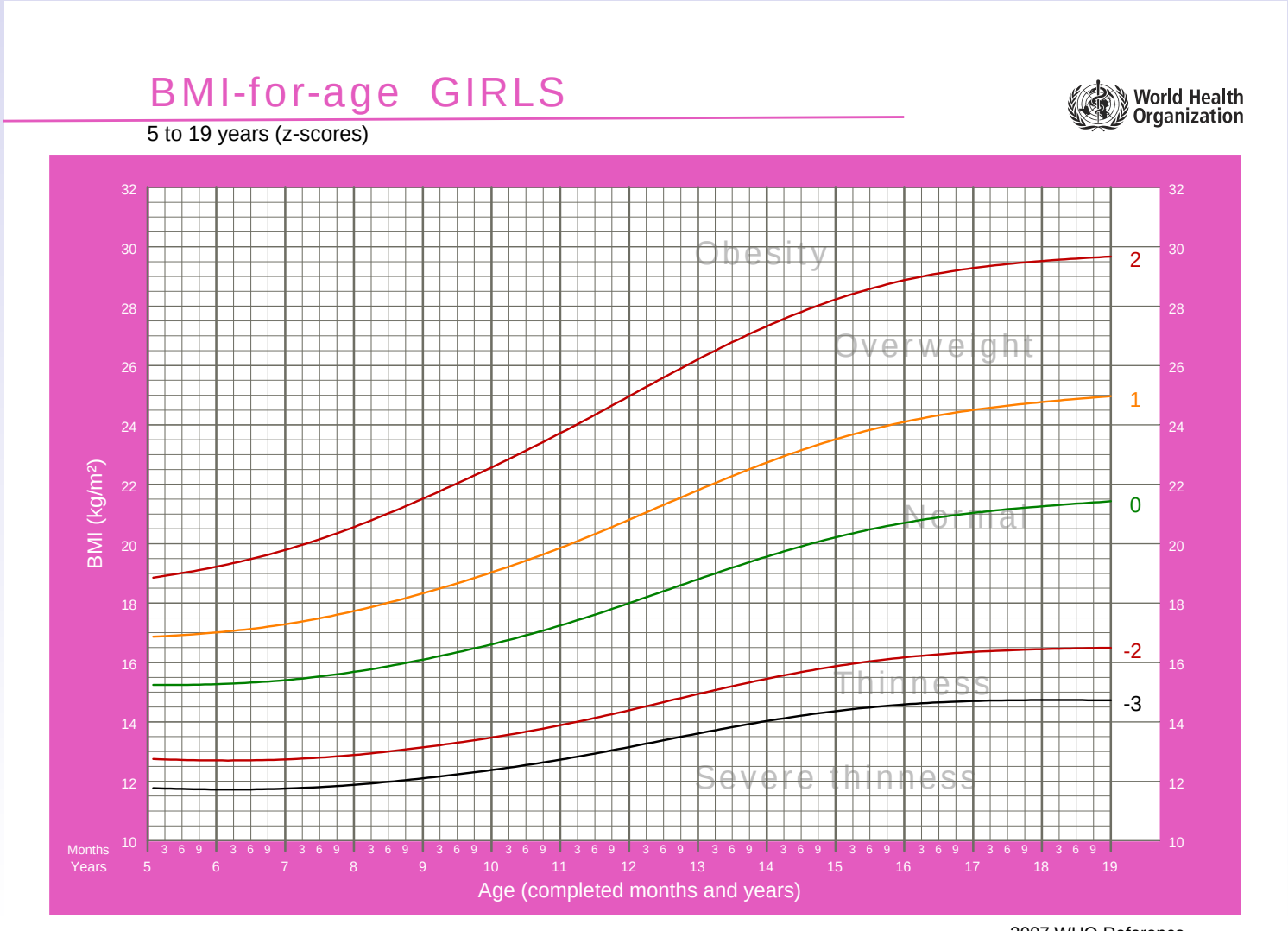
gamlss

The Dutch boys data: fan plot



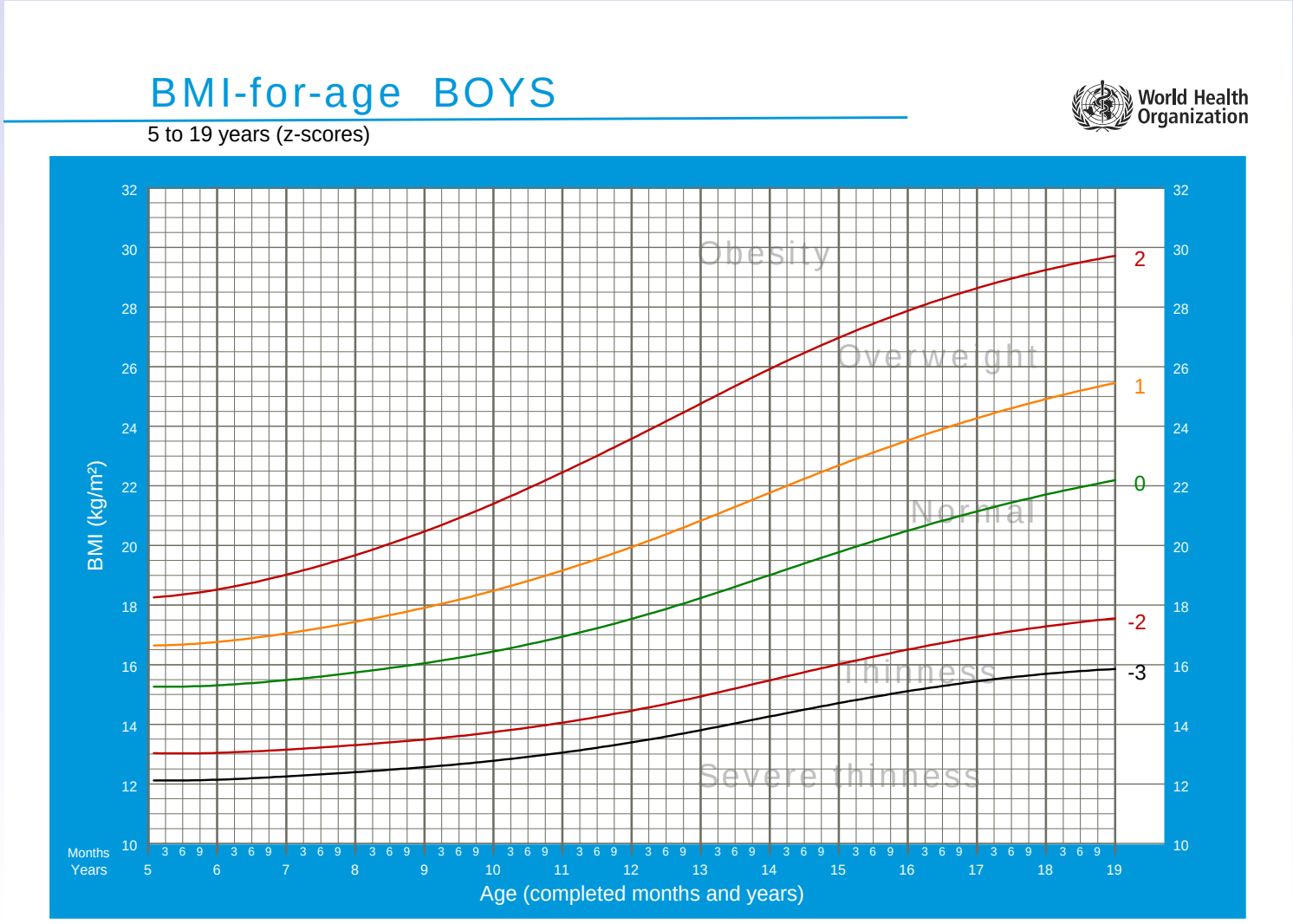
gamlss

World Health Organisation Child Growth Standards: Girls



gamlss

World Health Organisation Child Growth Standards: Boys



gamlss

The fish species data

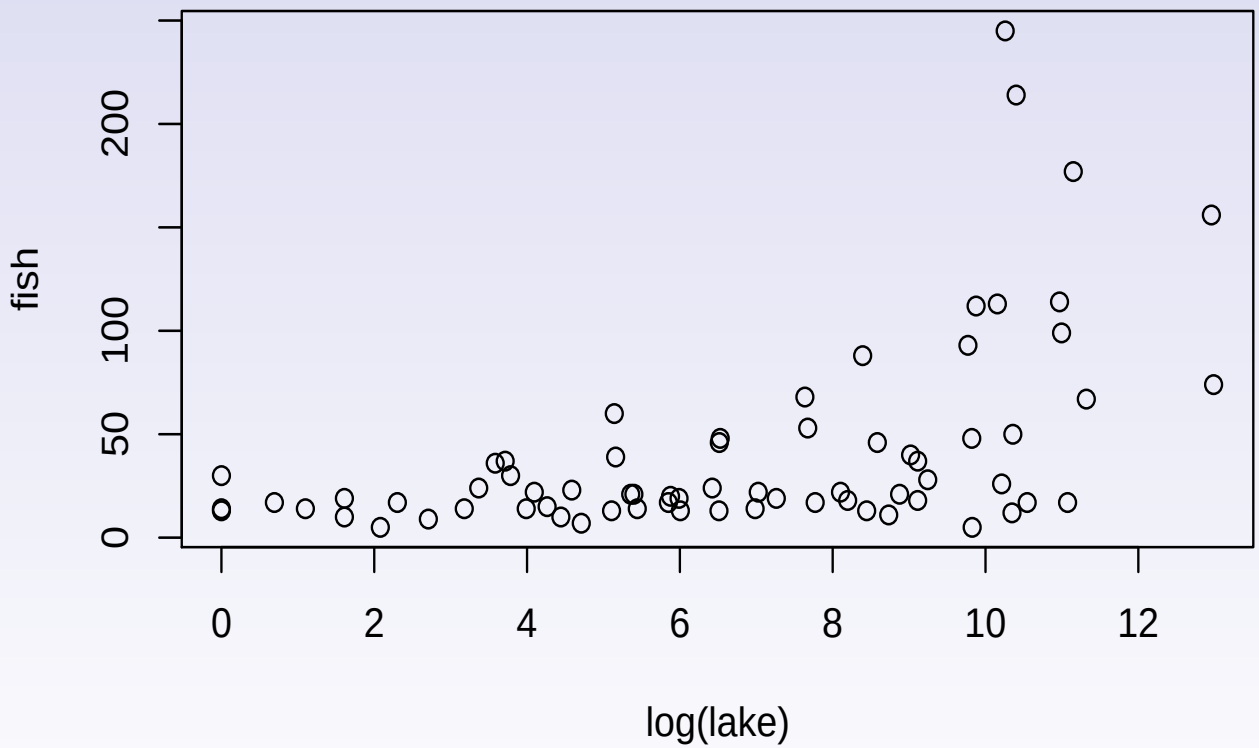
`fish` : the number of different species in 70 lakes in the world

`lake` : the lake area

Source: Stein and Juritz (1988)

gamlss

The fish species data



gamlss

A stylometric application

64 observations

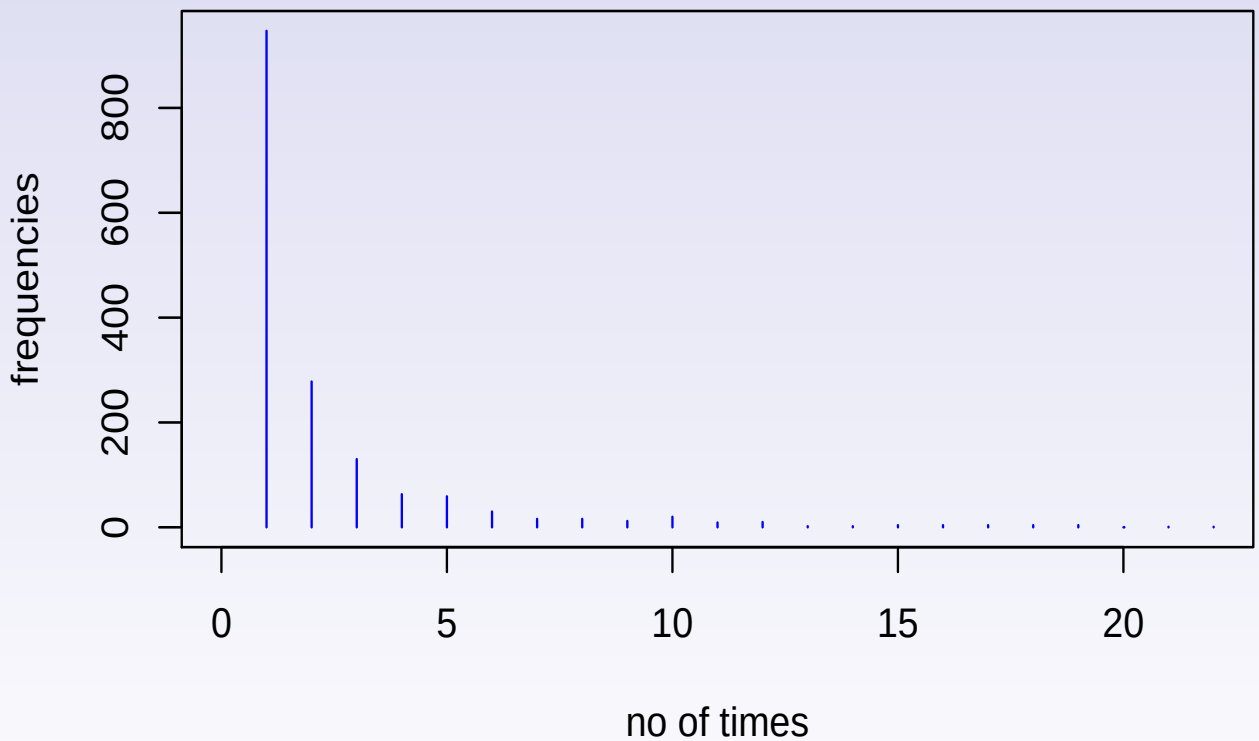
`word` : is the number of times a word appears in a single text

`freq` : the number of different words which occur exactly `word` times in the text

Source: Dr Mario Cortina-Borja

gamlss

The stylometric data



gamlss

The film data

4031 film openings in US 1988-1999:

`lborev1` : the response variable
log of the (total revenues - the first week revenue), calculated in 1987 prices

`lboopen` : the log the first week revenue, calculated in 1987 prices

`lnosc` : the log of the number of screens

`dist` : the type of distributor i.e. whether independent or major distributor

Source: Voudouris *et al.* (2010) Go to Quiz

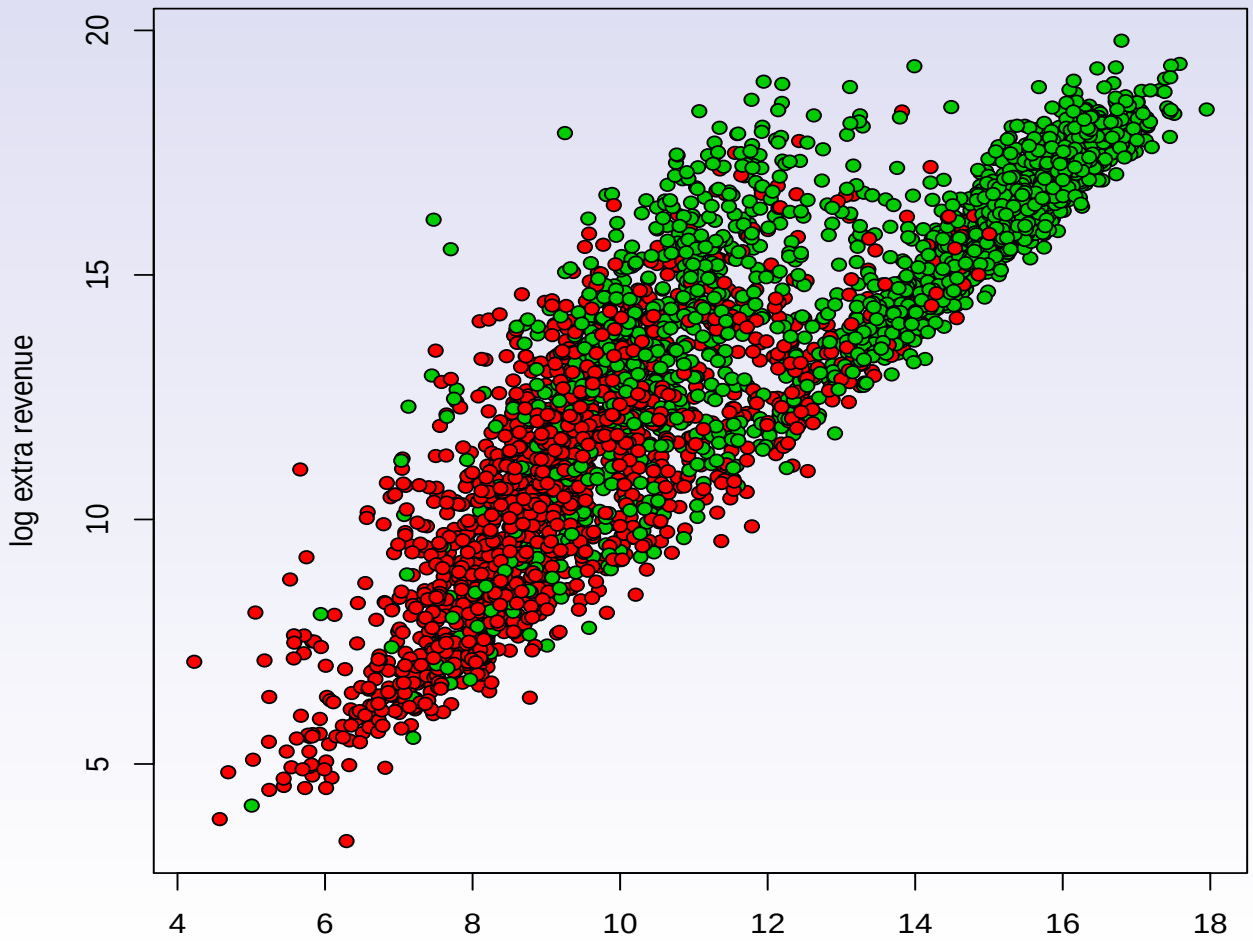
gamlss

The film data: Question of interest

- Checking whether the "nobody knows anything" principle is valid
- Can we predict the extra revenue given the revenue at the first week?
- Which other explanatory variables are important?
- What is the conditional distribution of the response variable given the explanatory variables?

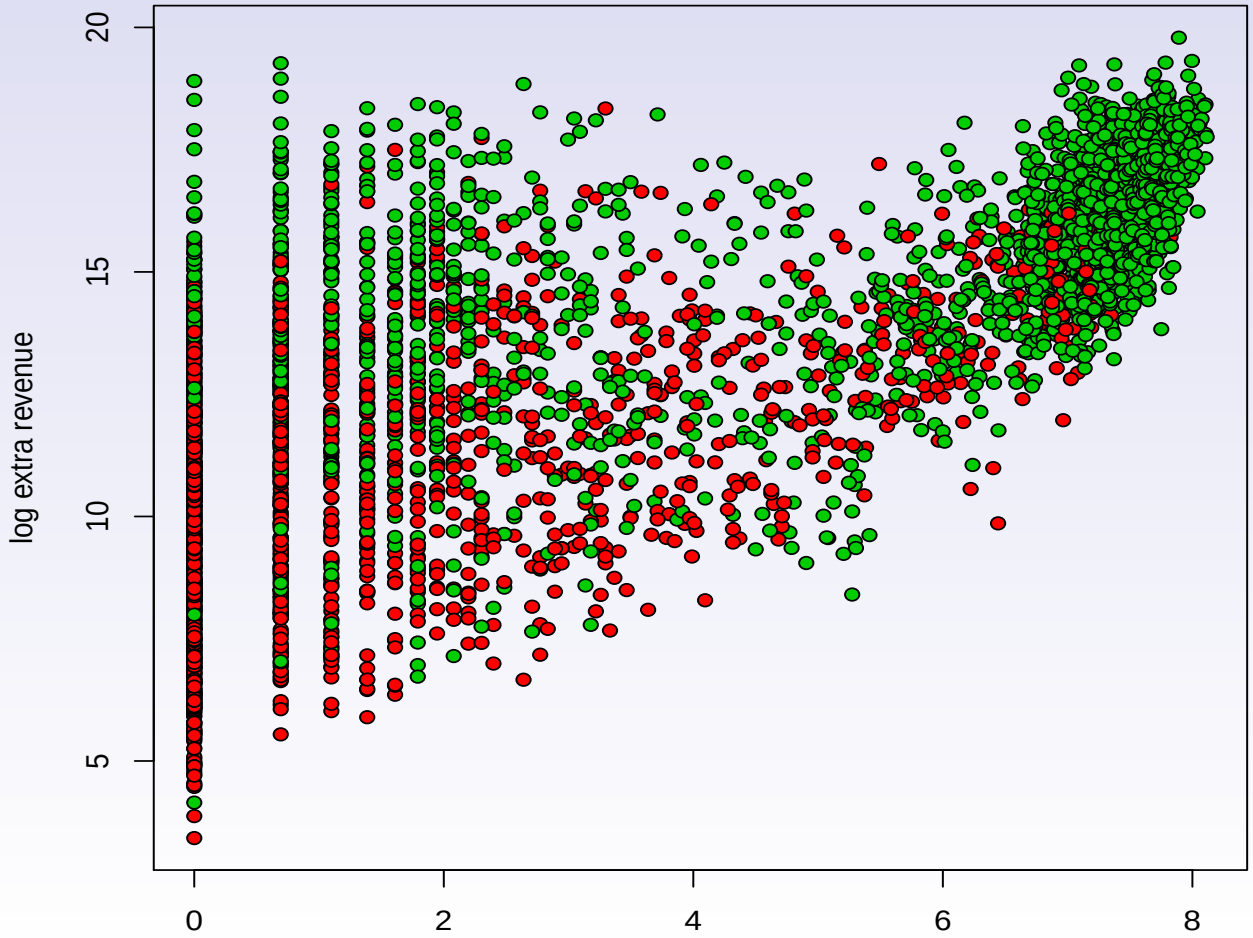
gamlss

The film data: log extra revenue against log first week revenue



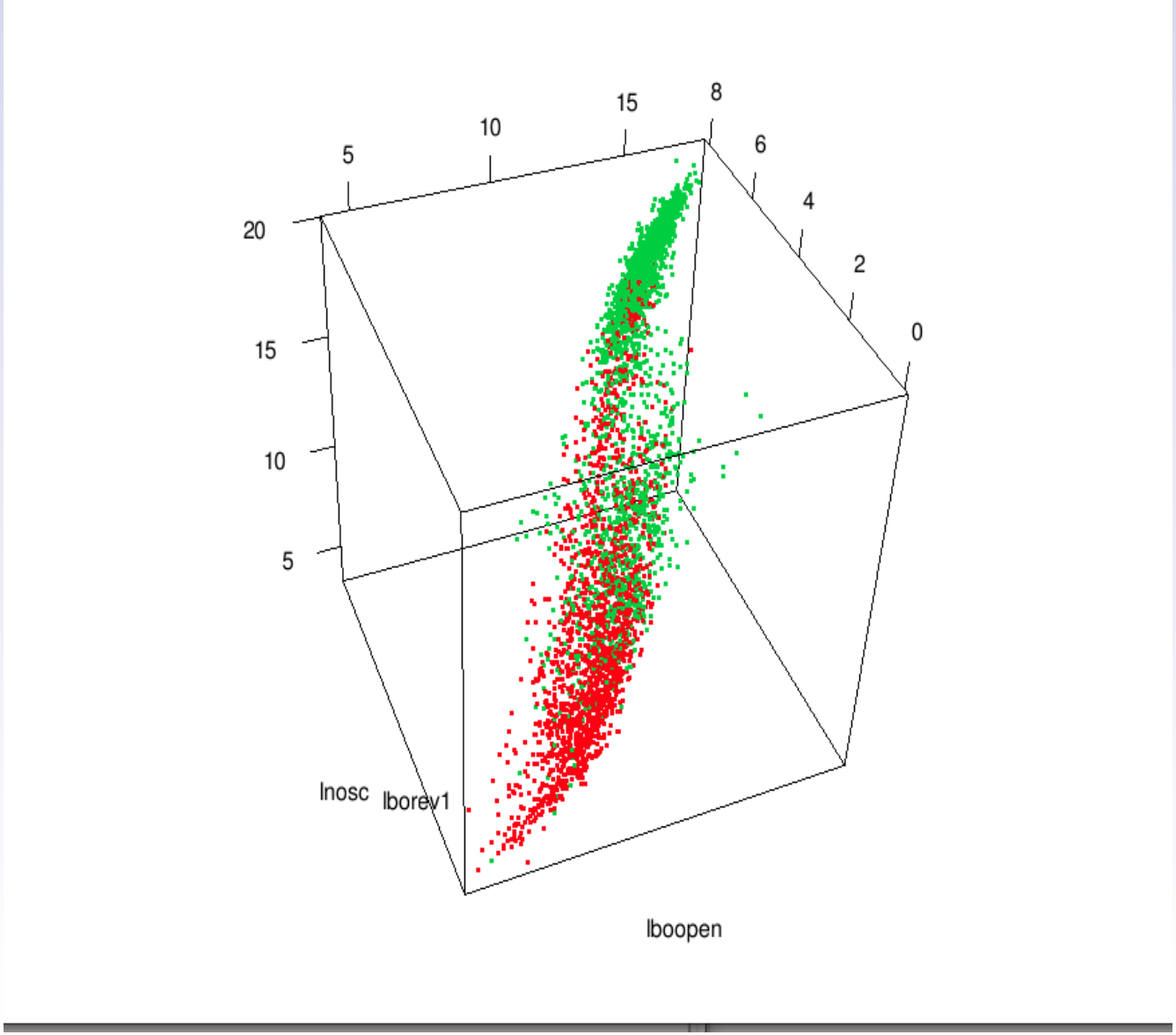
gamlss

The film data: log extra revenue against log number of screens



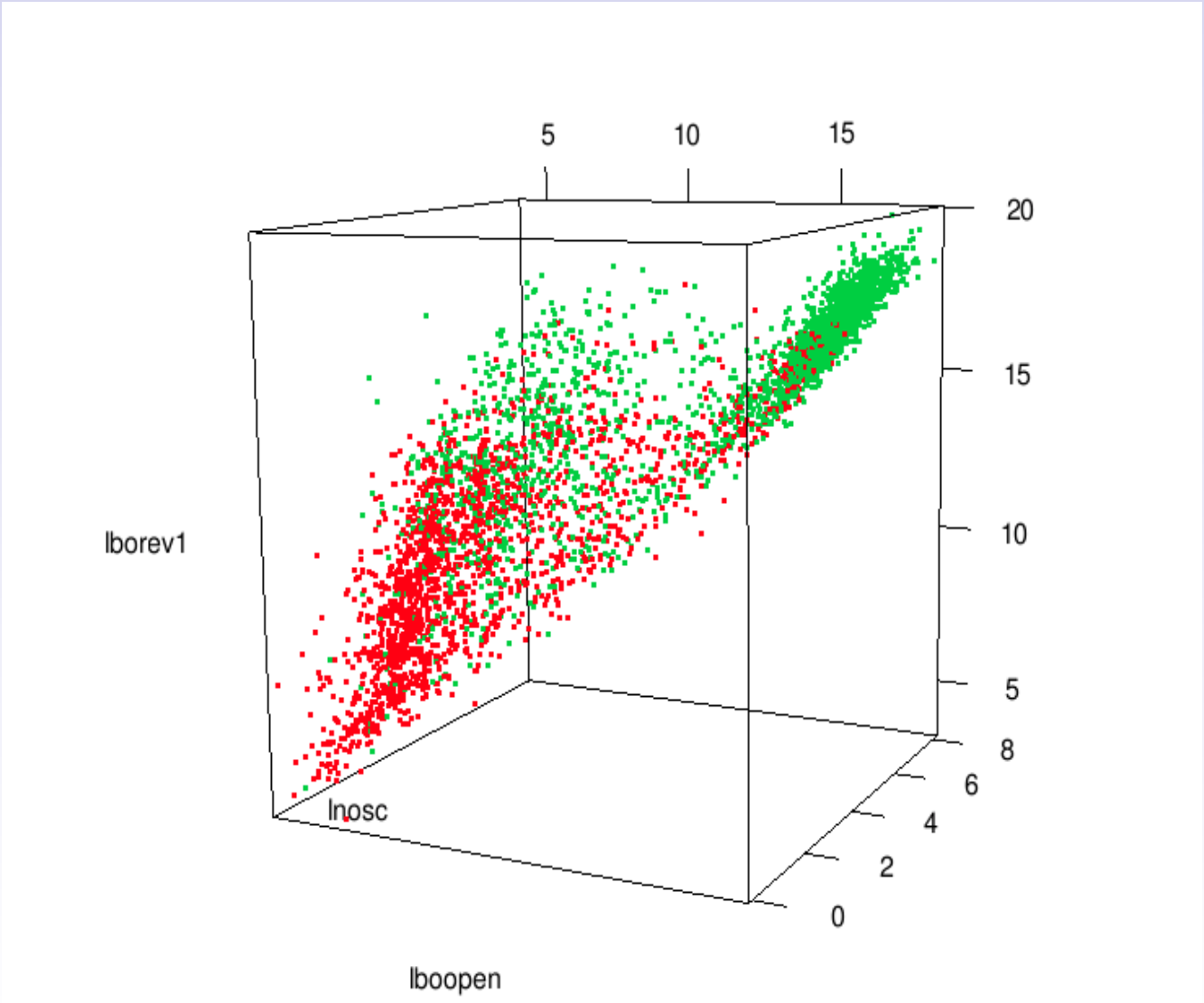
gamlss

The film data: 3d plots



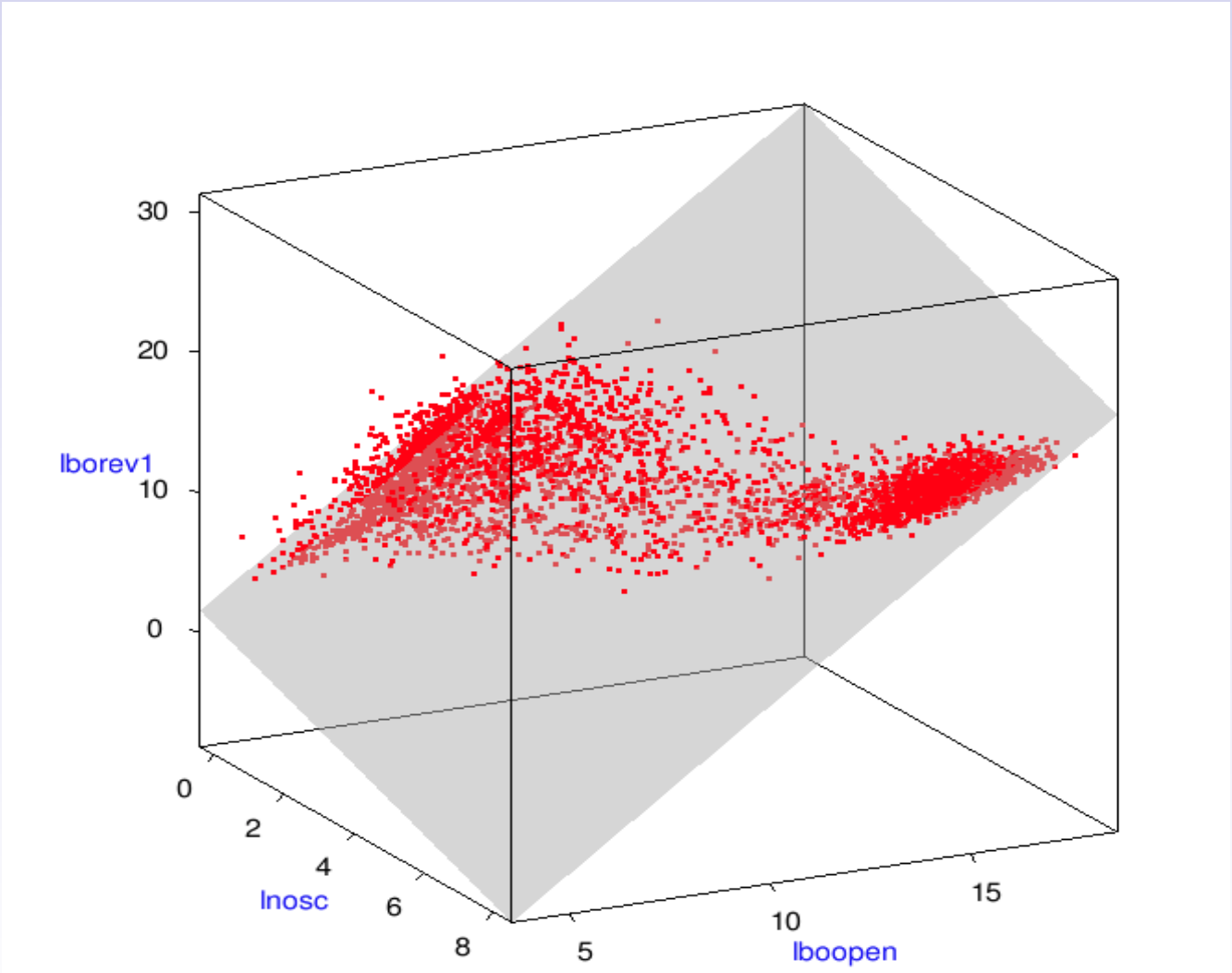
gamlss

The film data: 3d plots



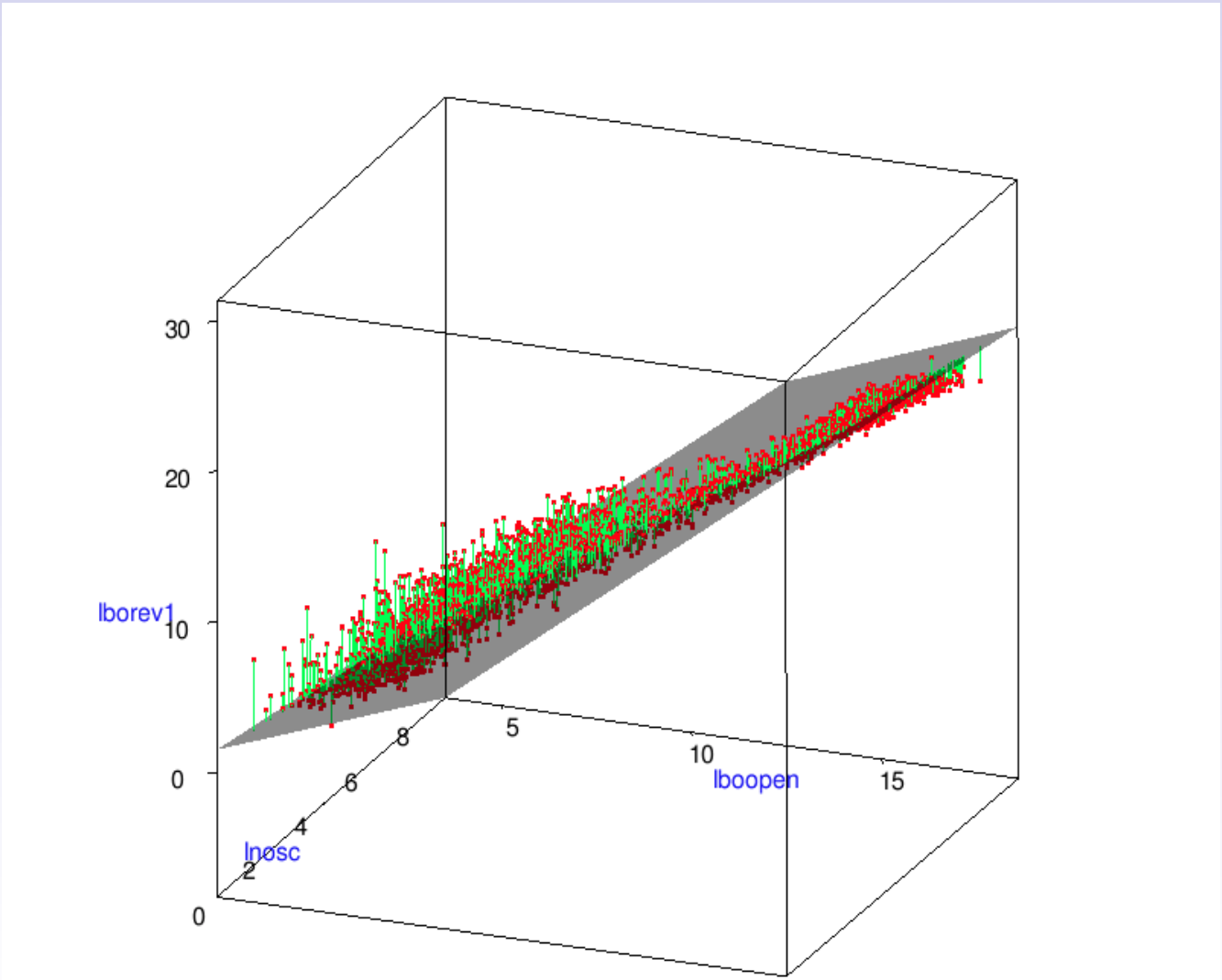
gamlss

The film data: Least Square fit



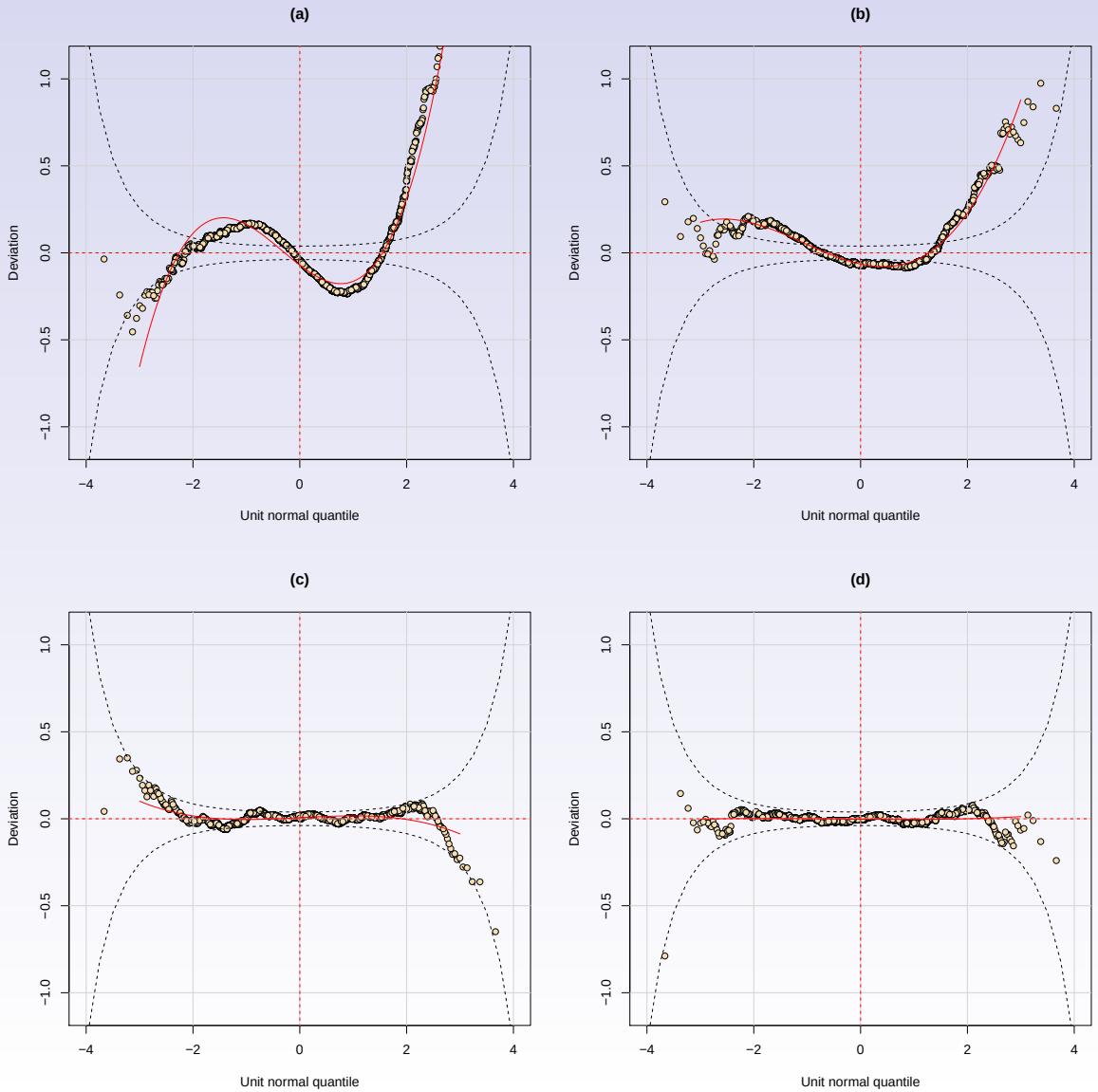
gamlss

The film data: the residuals from least squares fit



gamlss

The film data: worm plot of the residuals



gamlss

What we need for modelling the above data?

We need

- flexible distributions for the response variable
- to be able to deal with heterogeneity in the data
- to be able to model skewness and kurtosis
- to be able to model overdispersion in count data
- We need modelling all the parameters of the distributions
- flexible functions to model the relationship between the parameter of the distribution and the explanatory variables

gamlss

The statistical modelling philosophy

Statistical modelling is the art of building parsimonious statistical models for a better understanding of the phenomena of interest.

- Any model is a simplification of reality therefore no model is correct but some of them are useful
- Occam's Razor which states *'entities should not be multiplied beyond necessity'* or KISS (Keep It Simple Stupid)

gamlss

The statistical modelling philosophy

- Far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise. – John W. Tukey (1962), "The future of data analysis", Annals of Mathematical Statistics 33, 1-67 (page 13)

- *"no matter how beautiful your theory, no matter how clever you are or what your name is, if it disagrees with experiment, it's wrong"* (Richard Feynman)

- As statisticians we should try different models and choose the one appropriate for the data

gamlss

What is GAMLSS?

GAMLSS: are semi-parametric regression type models.

- regression type: we have many explanatory variables $\mathbf{X}$ and one response variable $\mathbf{y}$ and we believe that $\mathbf{X} \rightarrow \mathbf{y}$
- parametric: a parametric distribution assumption for the response variable,
- semi: the parameters of the distribution, as functions of explanatory variables, may involve non-parametric smoothing functions
- GAMLSS philosophy: try different models

GAMLSS is a generalization of GLM and GAM models.

gamlss

Regression type models

$$\texttt{input} \longrightarrow \texttt{output} \tag{1}$$

$$X \longrightarrow Y \tag{2}$$

X : *input* variable(s), *explanatory*, *independent* variable(s)

Y : *response* variable , *dependent* variable, *target*, *y-variable*, *output*

one response variable is *univariate* analysis otherwise *multivariate*

gamlss

The linear model

Linear regression model

$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ where $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \sigma^2)$

$\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$ where $\mathbf{y} \sim N(\boldsymbol{\mu}, \sigma^2)$

Assumptions

1. The X's effect Y linearly
2. The error terms are independently distributed random variables
3. The variance of $\boldsymbol{\epsilon}$ and therefor Y is constant

gamlss

The linear model assumptions

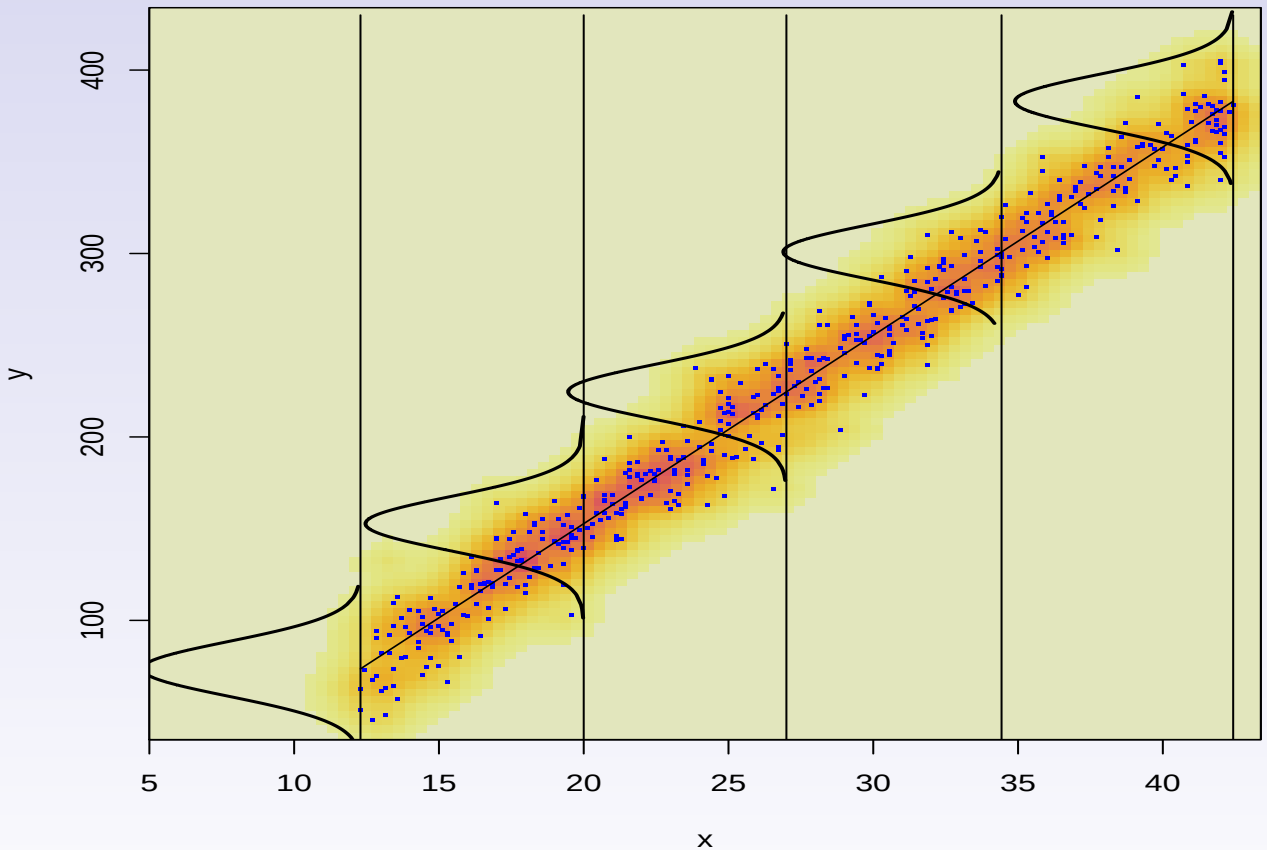


Figure: The simple regression assumptions

gamlss

The estimation of the linear model

Least Squares:

$$\boldsymbol{\epsilon}^\top \boldsymbol{\epsilon} = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^\top (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) = \sum_{i=1}^{n} \epsilon_i^2.$$

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{Y}$$

Maximum Likelihood

$$\ell\left(\boldsymbol{\beta}, \sigma^2\right) = \frac{n}{2} log\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2}\left(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}\right)^\top \left(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}\right)$$

$$\hat{\sigma}^2 = \frac{\left(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}\right)^\top \left(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}\right)}{n}.$$

gamlss

The weighted linear model

Weighted Linear regression model

$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ where $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$

$\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$ where $\mathbf{y} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

often $\boldsymbol{\Sigma} = \sigma^2\mathbf{W}^{-1}$
where $\mathbf{W}$ is a diagonal matrix

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}^\top\mathbf{W}\mathbf{X}\right)^{-1}\mathbf{X}^\top\mathbf{W}\mathbf{y}. \tag{3}$$

gamlss

The weighted linear model assumptions

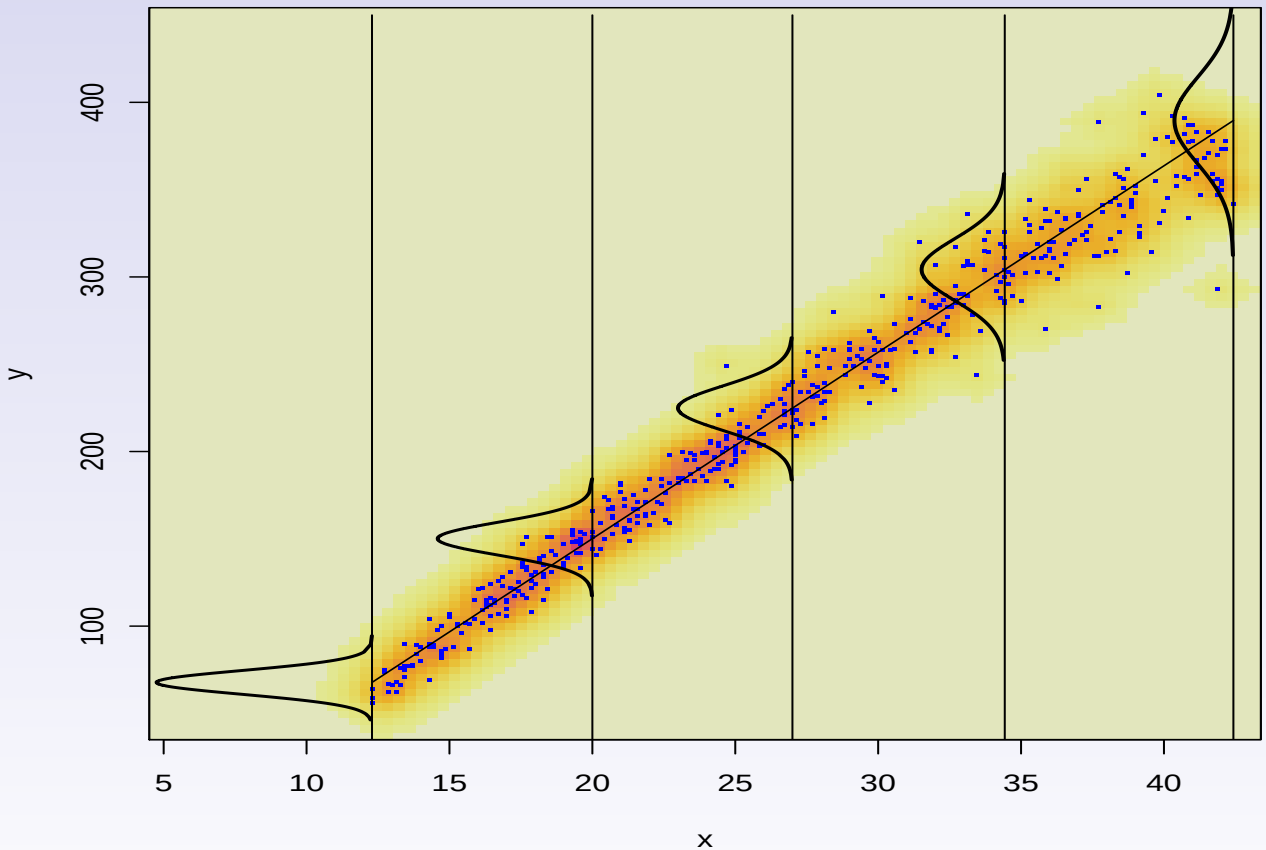


Figure: The weighted least squares regression assumptions

gamlss

The generalised linear model

generalised Linear Model, Wedderburn and Nelder (1972)

$g(\boldsymbol{\mu}) = \mathbf{X}\boldsymbol{\beta}$ where $\mathbf{y} \sim \mathit{ExpFamily}(\boldsymbol{\mu}, \phi)$

where $g()$ is the link function

The exponential family

1. normal
2. Gamma
3. inverse Gaussian
4. Poisson
5. binomial

gamlss

The generalised linear model

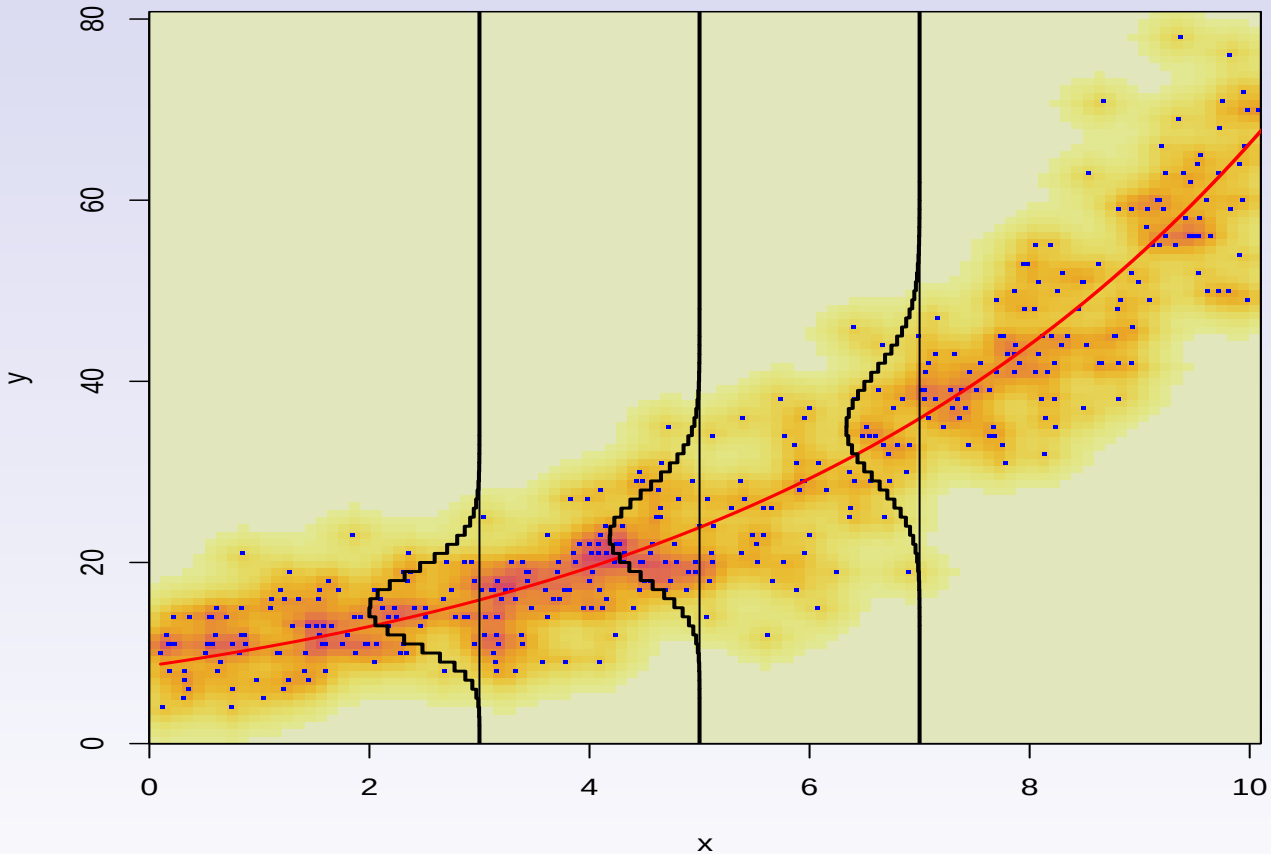


Figure: The Poisson regression assumptions

gamlss

The generalised linear model: parameter estimation

Maximize the log-Likelihood with respect to $\boldsymbol{\beta}$.

$$\ell(\boldsymbol{\theta}) = \log L(\boldsymbol{\theta}) = \sum_{i=1}^{n} \log f(y_i|\boldsymbol{\theta})$$

This is done using Iterative Weighted Least Squares

$$\hat{\boldsymbol{\beta}}^{(i)} = \left(\mathbf{X}^\top \mathbf{W}^{(\mathbf{i})} \mathbf{X}\right)^{-1} \mathbf{X}^\top \mathbf{W}^{(\mathbf{i})} \mathbf{z}^{(\mathbf{i})}. \tag{4}$$

where $\mathbf{z}^{(i)}$ iterative response variable and $\mathbf{W}^{(\mathbf{i})}$ are the iterative weights.

gamlss

The generalised additive model

generalised additive model Hastie and Tibshirani (1990)

$g(\boldsymbol{\mu}) = \mathbf{X}\boldsymbol{\beta} + h_1(\mathbf{x}_1) + ... + h_k(\mathbf{x}_k)$ where $\mathbf{y} \sim ExpFamily(\boldsymbol{\mu}, \phi)$

where $h_j(\mathbf{x}_j)$ are smooth functions of the X's.
To estimate the parameters $\boldsymbol{\beta}$ and the smoothing functions maximize the penalised log-Likelihood

$$\ell_p = \ell - \frac{1}{2}\sum_{j=1}^{J_k} \lambda_j \mathbf{Q}_j(\mathbf{h}_j)$$

where $\mathbf{Q}_j(\mathbf{h}_j)$ is a quadratic function of the h_j's

gamlss

The generalised additive model

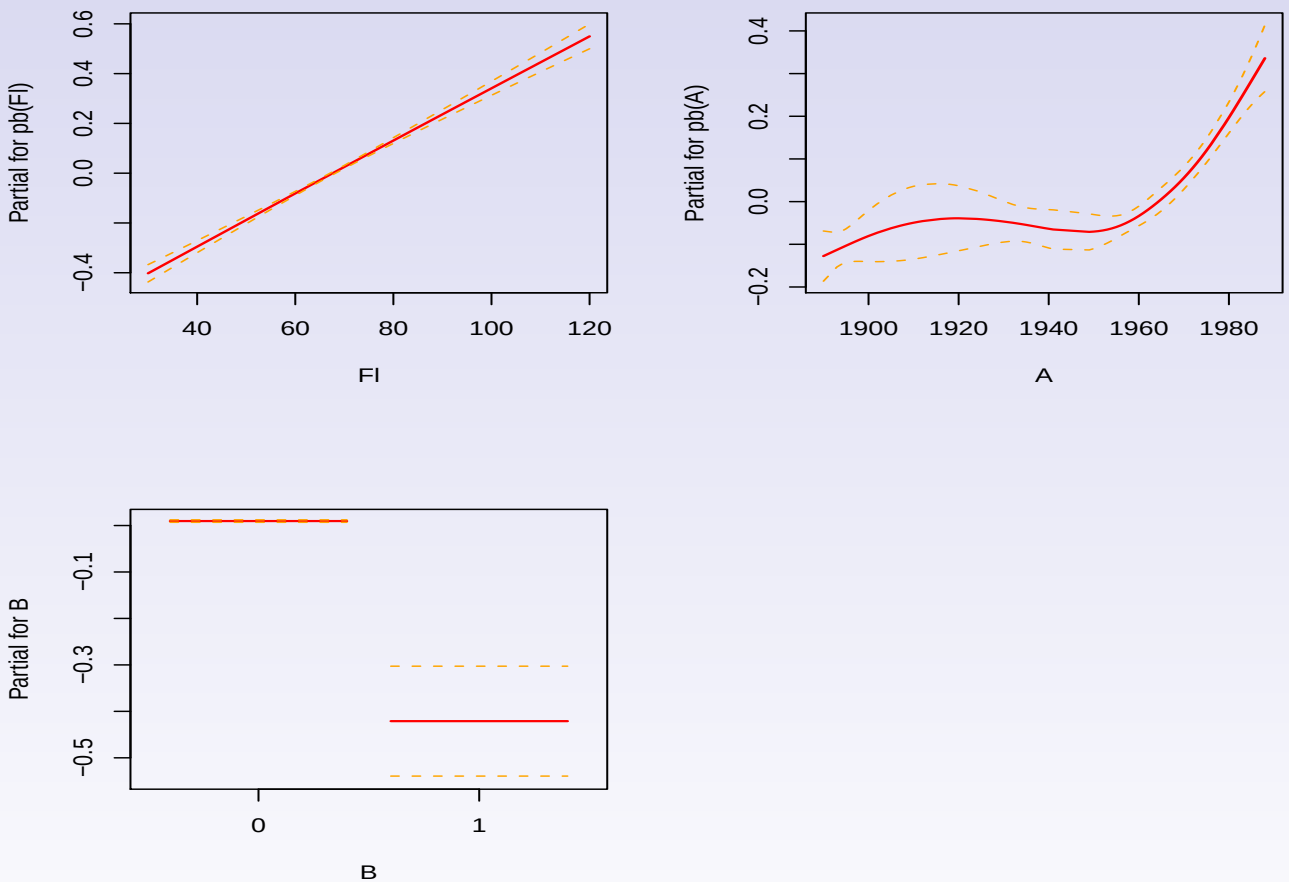
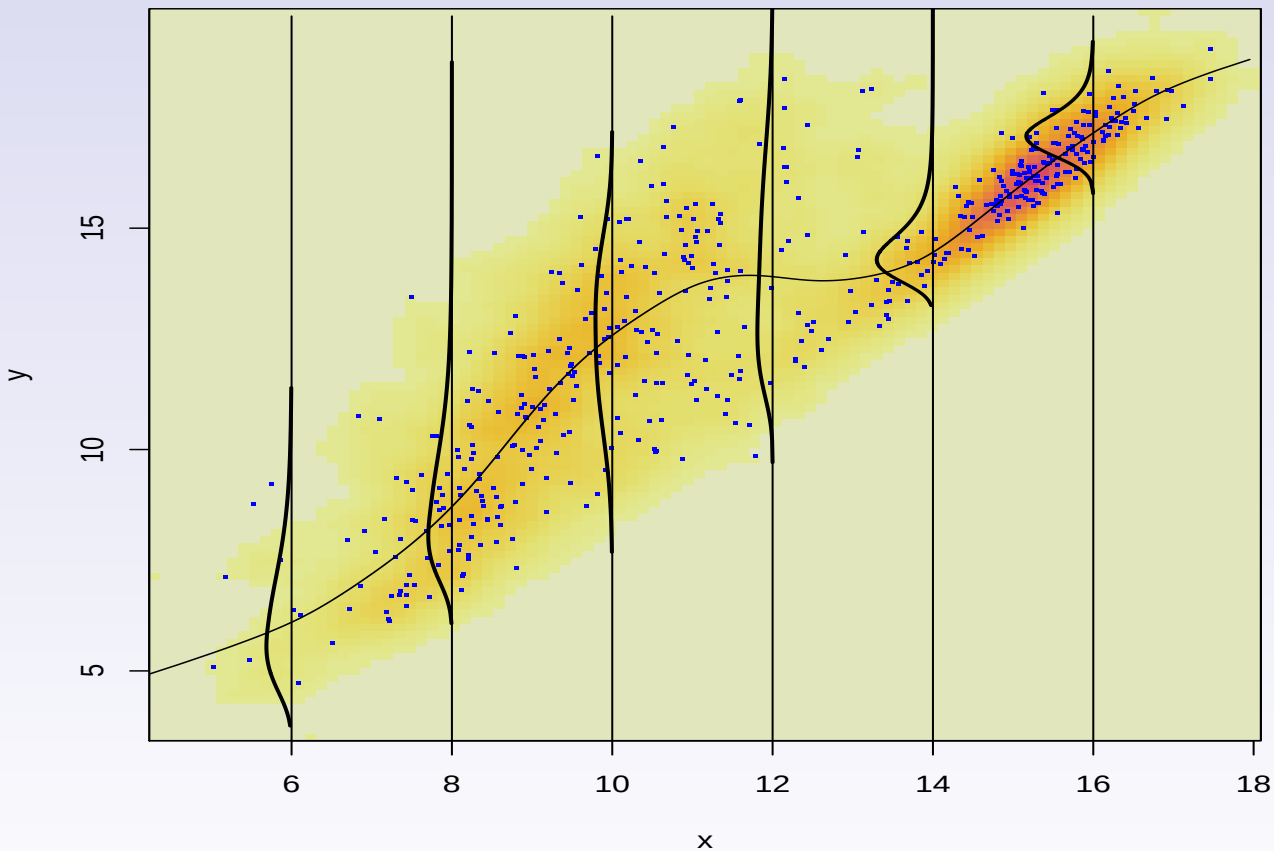


Figure: A generalised additive model fit

gamlss

GAMLSS: the model



gamlss

GAMLSS: the model

Let Y is the response variable

$$Y_i \sim f_Y(y_i|\theta_1, \theta_2, \theta_3, \theta_4) = f_Y(y_i|\mu_i, \sigma_i, \nu_i, \tau_i)$$

where $f_Y()$ can be any distribution (non only the exponential family).
any parameter θ_i can be a function of explanatory variables.

gamlss

GAMLSS: the model

$$
\begin{aligned}
g_1(\boldsymbol{\mu}) &= \boldsymbol{\eta}_1 = \mathbf{X}_1\boldsymbol{\beta}_1 + \sum_{j=1}^{J_1} h_{j1}(x_{j1}) \\
g_2(\boldsymbol{\sigma}) &= \boldsymbol{\eta}_2 = \mathbf{X}_2\boldsymbol{\beta}_2 + \sum_{j=1}^{J_2} h_{j2}(x_{j2}) \\
g_3(\boldsymbol{\nu}) &= \boldsymbol{\eta}_3 = \mathbf{X}_3\boldsymbol{\beta}_3 + \sum_{j=1}^{J_3} h_{j3}(x_{j3}) \\
g_4(\boldsymbol{\tau}) &= \boldsymbol{\eta}_4 = \mathbf{X}_4\boldsymbol{\beta}_4 + \sum_{j=1}^{J_4} h_{j4}(x_{j4}).
\end{aligned}
$$

gamlss

GAMLSS: the model (more general)

$$g_1(\boldsymbol{\mu}) = \boldsymbol{\eta}_1 = \mathbf{X}_1\boldsymbol{\beta}_1 + \sum_{j=1}^{J_1}\mathbf{Z}_{j1}\boldsymbol{\gamma}_{j1}$$

$$g_2(\boldsymbol{\sigma}) = \boldsymbol{\eta}_2 = \mathbf{X}_2\boldsymbol{\beta}_2 + \sum_{j=1}^{J_2}\mathbf{Z}_{j2}\boldsymbol{\gamma}_{j2}$$

$$g_3(\boldsymbol{\nu}) = \boldsymbol{\eta}_3 = \mathbf{X}_3\boldsymbol{\beta}_3 + \sum_{j=1}^{J_3}\mathbf{Z}_{j3}\boldsymbol{\gamma}_{j3}$$

$$g_4(\boldsymbol{\tau}) = \boldsymbol{\eta}_4 = \mathbf{X}_4\boldsymbol{\beta}_4 + \sum_{j=1}^{J_4}\mathbf{Z}_{j4}\boldsymbol{\gamma}_{j4}.$$

gamlss

GAMLSS: the random effects

where each γ_{jk} is a random variable

$$\gamma_{jk} \sim N_{q_{jk}}(\mathbf{0}, \mathbf{G}_{jk}^{-1})$$

where $\mathbf{G}_{jk}^{-1}$ is the (generalised) inverse of a symmetric matrix

$$\mathbf{G}_{jk} = \mathbf{G}_{jk}(\boldsymbol{\lambda}_{jk})$$

hyperparameters: $\boldsymbol{\lambda}_{jk}$

gamlss

GAMLSS: submodels

semi-parametric additive

$$g_k(\boldsymbol{\theta}_k) = \boldsymbol{\eta}_k = \mathbf{X}_k \boldsymbol{\beta}_k + \sum_{j=1}^{J_k} h_{jk}(\mathbf{x}_{jk})$$

parametric linear

$$g_k(\boldsymbol{\theta}_k) = \boldsymbol{\eta}_k = \mathbf{X}_k \boldsymbol{\beta}_k$$

non-linear parametric + additive

$$g_k(\boldsymbol{\theta}_k) = \boldsymbol{\eta}_k = h_k(\mathbf{X}_k, \boldsymbol{\beta}_k) + \sum_{j=1}^{J_k} h_{jk}(\mathbf{x}_{jk})$$

gamlss

GAMLSS: Distributions

There are around 70 discrete discrete, continuous continuous, and mixed distributions, mixed, implemented as `gamlss.family` in the R including highly skew and kurtotic distributions Shapes,

- creating a new distribution is relatively easy
- truncating truncated an existing distribution
- using a censored version of an existing distribution
- mixing mixture different distributions to create a new finite mixture distribution.
- discretise discretise continuous distributions
- log or logic any continuous distribution in $(-\infty, \infty)$

next page Additive

gamlss

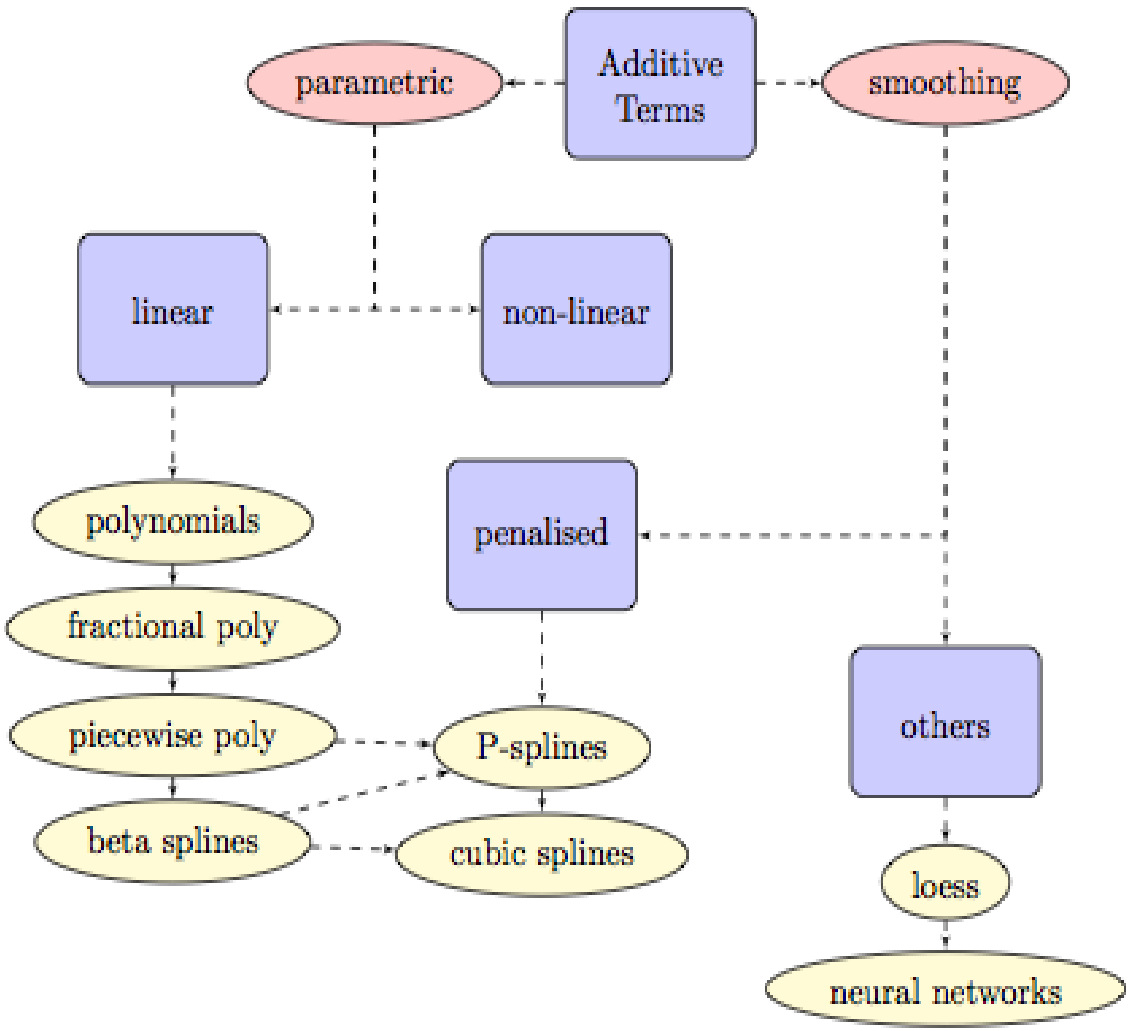
GAMLSS: additive terms

Additive terms	R function names
boosting	`boost()`
cubic splines based	`cs()`, `scs()`
decision trees	`tr()`
fractional and power polynomials	`fp()`, `pp()`
free knot smoothing (break points)	`fk()`
Penalised lgs	`la()`
`loess`	`lo()`
neural networks	`nn()`
non-linear fit	`nl()`
penalized Beta splines based	`pb()`, `ps()`, `cy()`, `pvc()`
random effects	`random()`
ridge regression	`ri()`, `ridge()`
Simon Wood's gam	`ga()`

gamlss

Table: Additive terms implemented within the **gamlss** packages

GAMLSS: additive terms



gamlss

What is GAMLSS?: The quantities involved

- y response variable
- $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_p)$ fixed effect design matrices
- $\mathbf{Z} = (\mathbf{Z}_1, \mathbf{Z}_2, \ldots, \mathbf{Z}_p)$ random effect design matrices
- $\boldsymbol{\beta}^\top = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top, \ldots, \boldsymbol{\beta}_p^\top)$ fixed effect parameters
- $\boldsymbol{\gamma}^\top = (\boldsymbol{\gamma}_1^\top, \boldsymbol{\gamma}_2^\top, \ldots, \boldsymbol{\gamma}_p^\top)$ random effect parameters
- $\boldsymbol{\lambda}^\top = (\lambda_{11}^\top, \lambda_{21}^\top, \ldots, \boldsymbol{\lambda}_{J_p p}^\top)$ hyperparameters

gamlss

Model estimation

Model estimation in GAMLSS

case 1 (Linear or non-linear parametric) GAMLSS model i.e $\boldsymbol{\beta}$'s but no $\boldsymbol{\gamma}$'s or $\boldsymbol{\lambda}$'s

case 2 Random effects or semi-parametric GAMLSS model i.e $\boldsymbol{\beta}$'s, $\boldsymbol{\gamma}$'s and $\boldsymbol{\lambda}$'s

gamlss

Model estimation

In case 1 the parametric case: GAMLSS estimates the $\boldsymbol{\beta}$'s by

maximum likelihood estimation i.e. maximizing the log-likelihood of the data.

$$\ell = \sum_{i=1}^{n} \log f(y_i|\mu_i, \sigma_i, \nu_i, \tau_i)$$

with respect to $\boldsymbol{\beta}$'s.

gamlss

Model estimation

In case 2 the random effect or semi-parametric case:

The main GAMLSS algorithm estimates $(\boldsymbol{\beta}, \gamma)$'s for given $\boldsymbol{\lambda}$'s by maximizing the penalized log-likelihood (equivalent to posterior mode estimation (MAP) assuming a uniform prior for $\boldsymbol{\beta}$'s):

$$\ell_p = \ell - \frac{1}{2}\sum_{k=1}^{p}\sum_{j=1}^{J_k}\lambda_{jk}\boldsymbol{\gamma}'_{jk}\mathbf{G}_{jk}\boldsymbol{\gamma}_{jk}$$

gamlss

Model estimation

MAP estimation of $(\boldsymbol{\beta}, \boldsymbol{\gamma})$ given $\boldsymbol{\lambda}$
MAP=maximum *a posteriori*= posterior mode

$$f(\boldsymbol{\beta}, \boldsymbol{\gamma}|\mathbf{y}, \boldsymbol{\lambda}) \propto f(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\gamma}) f(\boldsymbol{\gamma}|\boldsymbol{\lambda}))$$

assuming a uniform prior for $\boldsymbol{\beta}$

gamlss

Model estimation

$$\begin{aligned}
\log f(\boldsymbol{\beta}, \boldsymbol{\gamma}|\mathbf{y}, \boldsymbol{\lambda}) &= \log f(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\gamma}) + \log f(\boldsymbol{\gamma}|\boldsymbol{\lambda})) + c_1(\mathbf{y}, \boldsymbol{\lambda}) \\
&= \ell_h + c_1(\mathbf{y}, \boldsymbol{\lambda}) \\
&= \log f(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\gamma}) - \frac{1}{2}\sum_{k=1}^{p}\sum_{j=1}^{J_k} \lambda_{jk}\boldsymbol{\gamma}_{jk}'\mathbf{G}_{jk}\boldsymbol{\gamma}_{jk} + c_2(\mathbf{y}, \boldsymbol{\lambda}) \\
&= \ell_p + c_2(\mathbf{y}, \boldsymbol{\lambda})
\end{aligned}$$

since $\boldsymbol{\gamma}_{jk} \sim N(\mathbf{0}, \mathbf{G}_{jk}^{-1})$ and assuming a uniform prior for $\boldsymbol{\beta}$

gamlss

Model estimation

MAP estimation of $(\boldsymbol{\beta}, \gamma)$ given $\boldsymbol{\lambda}$

Hence given given $\boldsymbol{\lambda}$
MAP estimation of $(\boldsymbol{\beta}, \gamma)$
$\equiv$ maximizing ℓ_h , hierarchical log likelihood
$\equiv$ maximaizing ℓ_p penalized log likelihood
with respect to $(\boldsymbol{\beta}, \gamma)$

gamlss

Model estimation

Estimation of the hyper parameters $\boldsymbol{\lambda}$

- maximizing a profile GAIC over $\boldsymbol{\lambda}$
- maximizing the marginal likelihood of $\boldsymbol{\lambda}$

gamlss

Algorithms

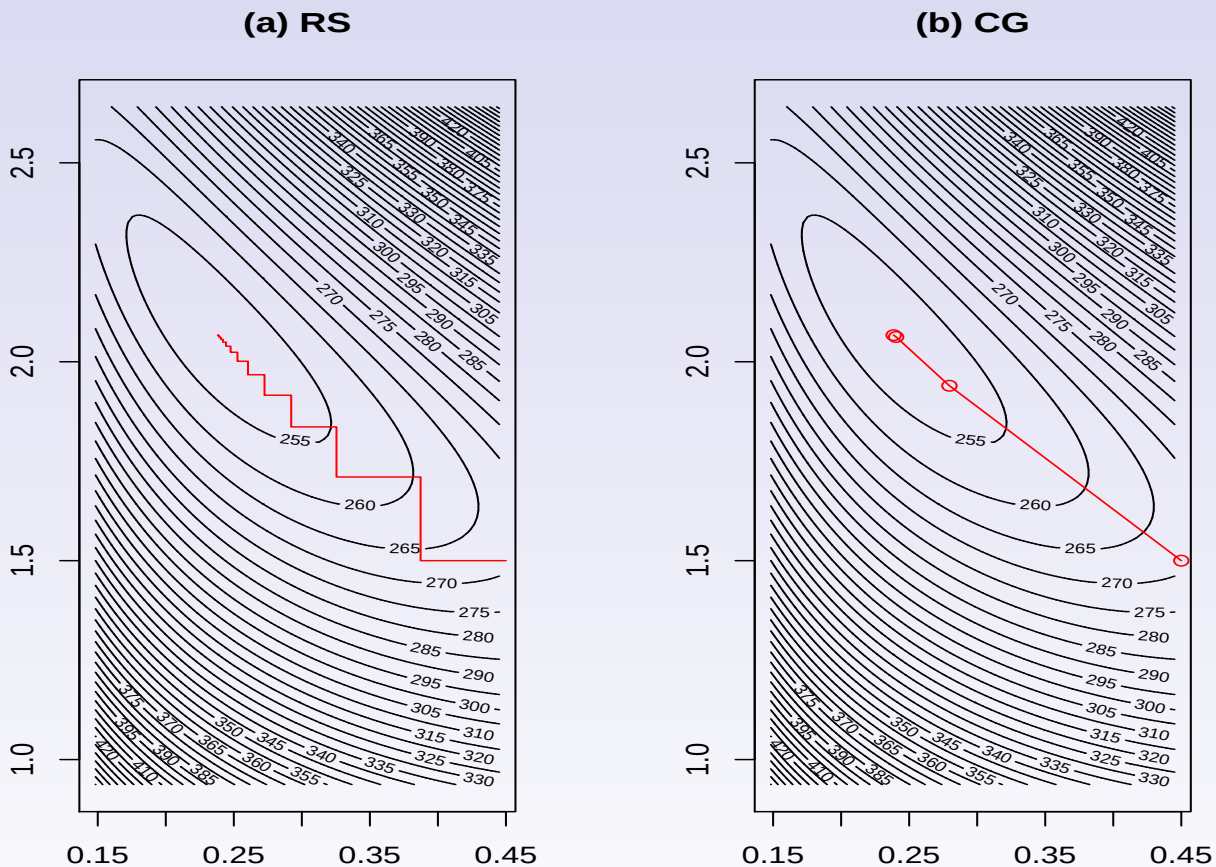
RS a generalization of the MADAM algorithm, Rigby and Stasinopoulos (1996a)

CG a generalization of Cole and Green (1992) algorithm

mixed a mixture of RS+CG (i.e. j iterations of RS, followed by k iterations of CG)

gamlss

Algorithms



Algorithms: properties

Advantages of algorithms

1. flexible modular fitting procedure
2. easy implementation of new distributions
3. easy implementation of new additive terms
4. simple starting values for (μ, σ, ν, τ) easily found
5. stable and reliable algorithms
6. fast fitting (for fixed hyperparameters)

gamlss

Algorithms: details

CG algorithm for MAP estimation of $(\boldsymbol{\beta}, \gamma)$
The r^{th} iteration of the Newton Raphson algorithm gives:

$$\begin{bmatrix} \mathbf{D}_{11} & \mathbf{D}_{12} \\ \mathbf{D}_{21} & \mathbf{D}_{22} \end{bmatrix}^{(r)} \left[\begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\gamma} \end{pmatrix}^{(r+1)} - \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\gamma} \end{pmatrix}^{(r)} \right] = \begin{bmatrix} \frac{\partial \ell_h}{\partial \boldsymbol{\beta}} \\ \frac{\partial \ell_h}{\partial \boldsymbol{\gamma}} \end{bmatrix}^{(r)}$$

where $\boldsymbol{\beta}^\top = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top, \ldots, \boldsymbol{\beta}_p^\top)$, $\boldsymbol{\gamma}^\top = (\boldsymbol{\gamma}_1^\top, \boldsymbol{\gamma}_2^\top, \ldots, \boldsymbol{\gamma}_p^\top)$ and ℓ_h is the hierarchical likelihood.
The Fisher scoring algorithm uses the expected information matrix instead of the observed information matrix $\mathbf{D}$

gamlss

Algorithms : Updating $\boldsymbol{\beta}_k$

Rearranging the row of the Fisher scoring equation for $\boldsymbol{\beta}_k$ gives

$$\mathbf{X}_k \boldsymbol{\beta}_k^{(r+1)} = \mathbf{H}_k^{(r)} \boldsymbol{\varepsilon}_{ok}^{(r)}$$

updates $\boldsymbol{\beta}_k$ where

$$\mathbf{H}_k^{(r)} = \mathbf{X}_k \left(\mathbf{X}_k^T \mathbf{W}_{kk}^{(r)} \mathbf{X}_k \right)^{-1} \mathbf{X}_k^T \mathbf{W}_{kk}^{(r)}$$

i.e. a weighted regression of partial residuals $\boldsymbol{\varepsilon}_{ok}^{(r)}$ against $\mathbf{X}_k$

gamlss

Algorithms : Updating γ_k

Rearranging the row of the Fisher scoring equation for γ_{jk} gives

$$\mathbf{Z}_{jk}\gamma_{jk}^{(r+1)} = \mathbf{S}_{jk}^{(r)}\varepsilon_{jk}^{(r)}$$

updates γ_{jk} where

$$\mathbf{S}_{jk}^{(r)} = \mathbf{Z}_{jk}\left(\mathbf{Z}_{jk}^{T}\mathbf{W}_{kk}^{(r)}\mathbf{Z}_{jk} + \mathbf{G}_{jk}\right)^{-1}\mathbf{Z}_{jk}^{T}\mathbf{W}_{kk}^{(r)}$$

i.e. a generalised ridge regression of partial residuals $\varepsilon_{jk}^{(r)}$ against $\mathbf{Z}_{jk}$

gamlss

Algorithms : convergence

Note that both the CG and the RS algorithms update each

$(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_p)$ and

$(\boldsymbol{\gamma}_{11}, \boldsymbol{\gamma}_{21}, \ldots, \boldsymbol{\gamma}_{J_p,p})$

until convergence of $(\boldsymbol{\beta}^{(r+1)}, \boldsymbol{\gamma}^{(r+1)})$

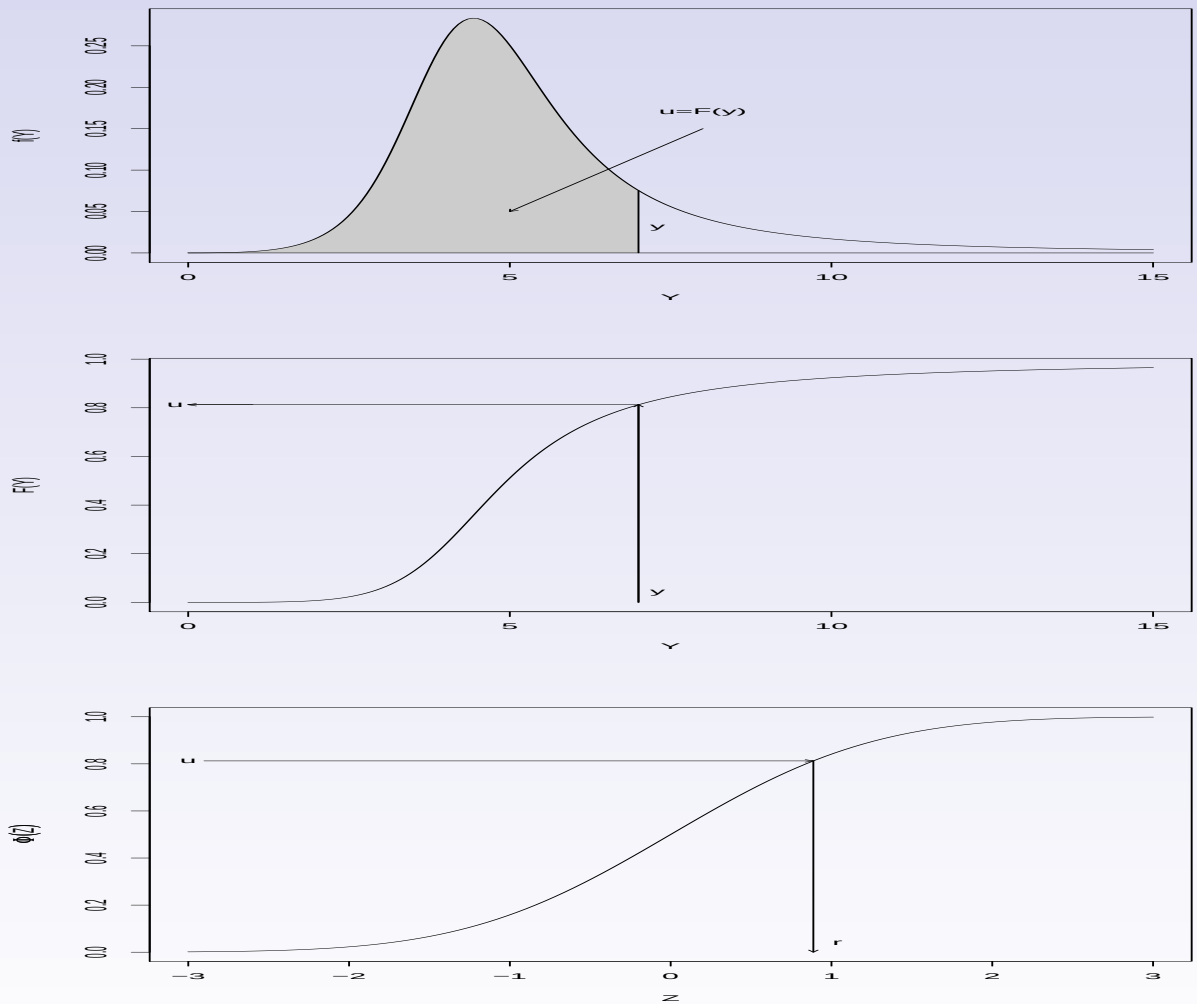
gamlss

Normalised quantile residuals-continuous response

If y_i is an observation from a continuous response variable then
$\hat{u}_i = F(y_i|\hat{\boldsymbol{\theta}}^i)$
where $u_i = F(y_i|\boldsymbol{\theta}^i)$ is the assumed cumulative distribution function for case i.
$r_i = \Phi^{-1}(u_i)$,
So ri will have an approximately standard normal distribution
r_i= z-scores

gamlss

Normalised quantile residuals-continuous response



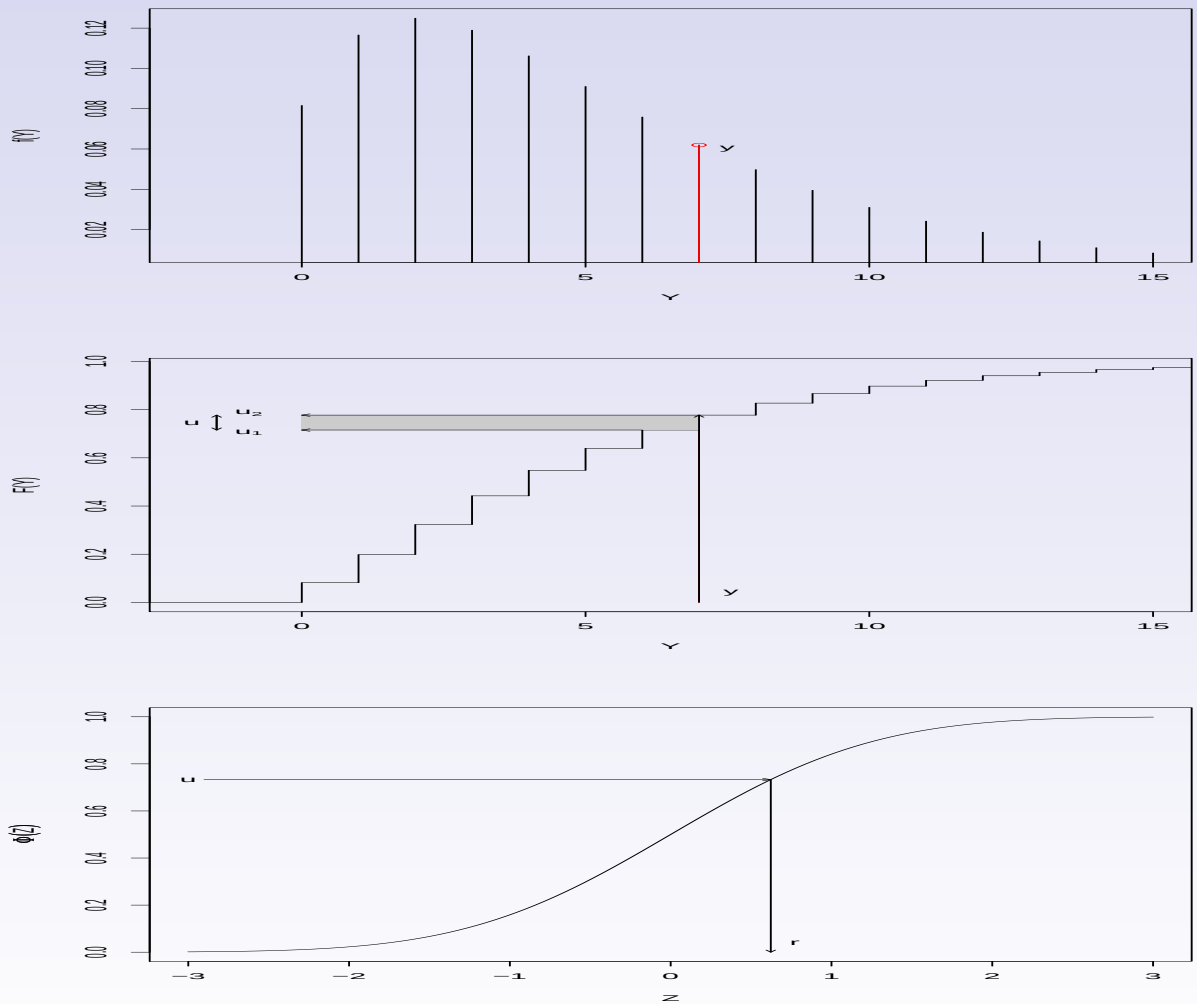
gamlss

Normalised quantile residuals-discrete response

If y_i is an observation from a discrete integer response then
$\hat{u}_i$ is a random value from the uniform distribution on the interval
$[u_1, u_2] = \left[F(y_i - 1|\hat{\boldsymbol{\theta}}^i), F(y_i|\hat{\boldsymbol{\theta}}^i)\right]$.
$\hat{r}_i = \Phi^{-1}(\hat{u}_i)$.

gamlss

Normalised quantile residuals-disctrete response



gamlss

Conclusions

- Unified framework for univariate regression type of models
- Allows any distribution for the response variable Y
- Models all the parameters of the distribution of Y
- Allows a variety of additive terms in the models for the distribution parameters
- The fitted algorithm is modular, where different components can be added easily
- Models can be fitted easily and fast
- Explanatory tool to find appropriate set of models
- It deals with overdispersion, skewness and kurtosis

gamlss

End

END

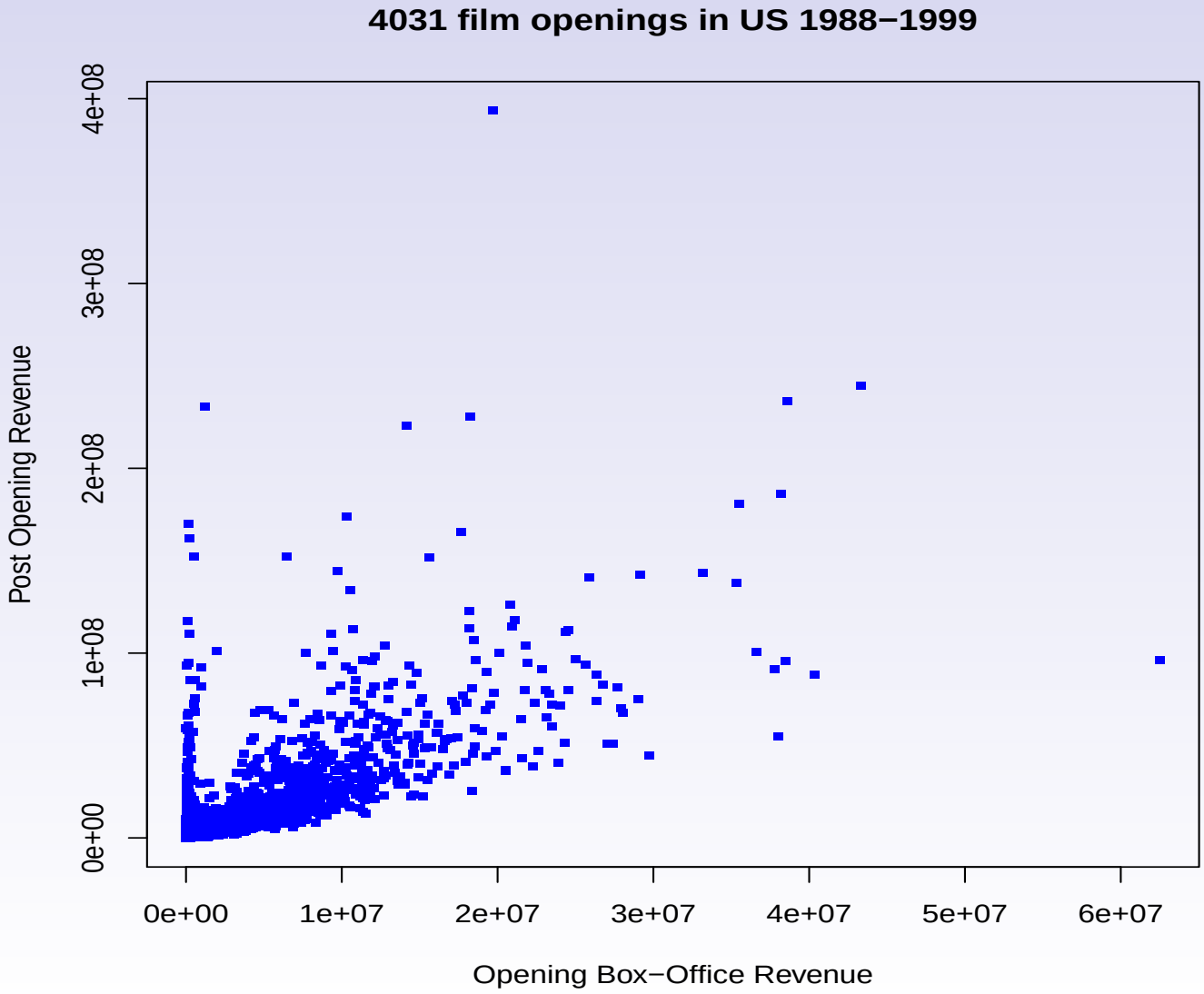
for more information see

www.gamlss.org

gamlss

Film Data

Go back



gamlss

Film Data: the quiz

A small Quiz

Question 1 which film in the 90's has made the most money?

Question 2 which film in the 90's did badly in the first week but subsequently did very well?

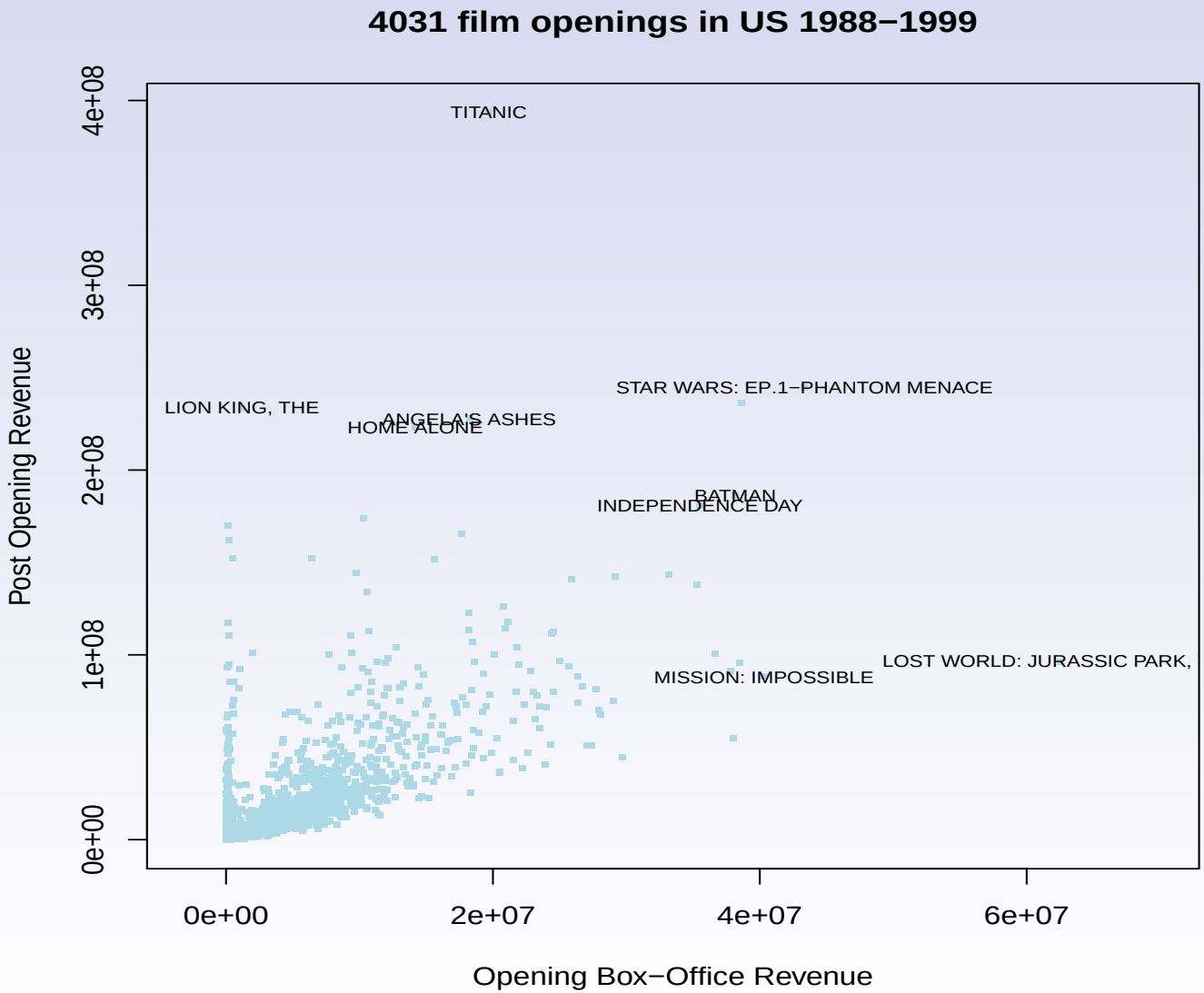
Question 3 which film in the 90's did well in the first week but flopped after that?

Go back

gamlss

Film Data: Answers to the quiz

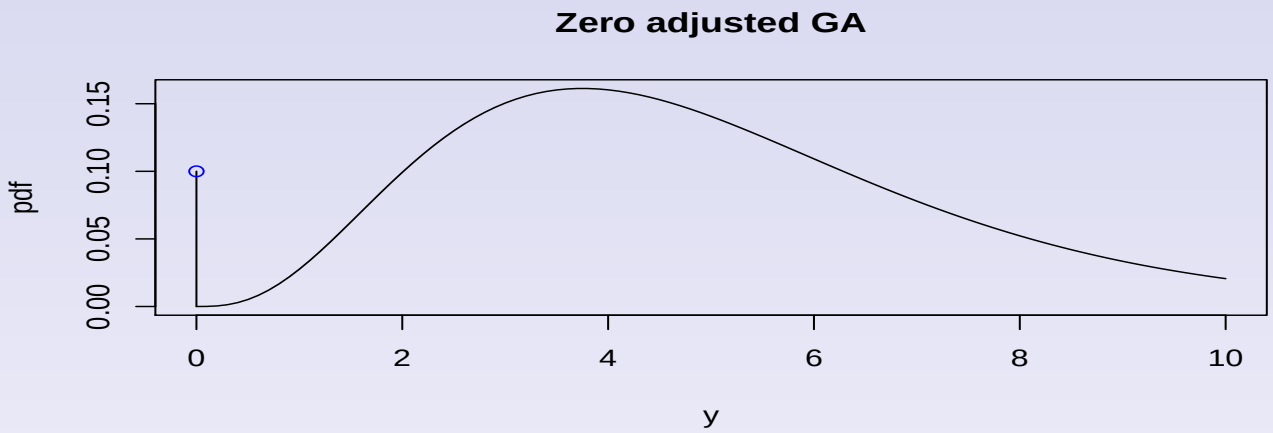
Go back

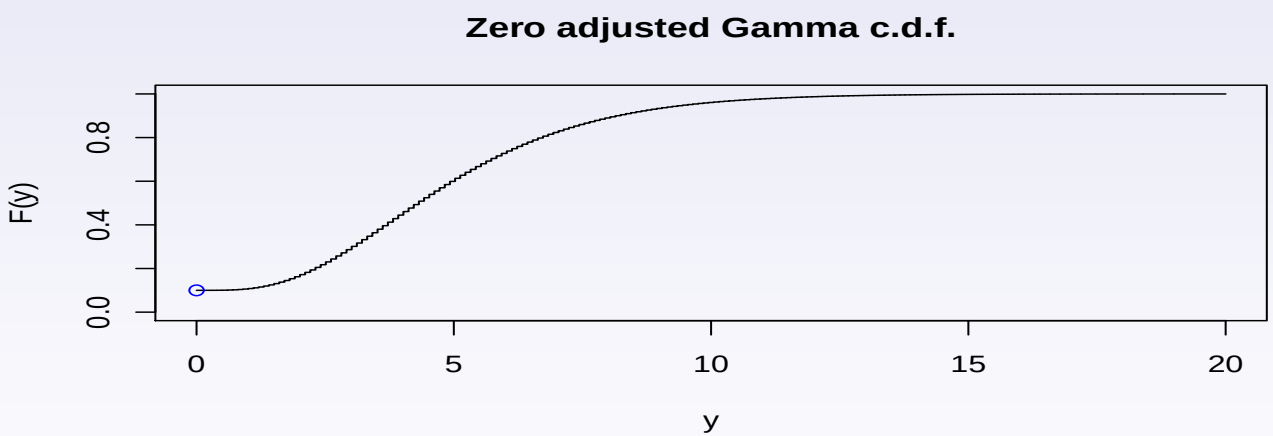


gamlss

Example of mixed distribution distributions

Go back Distributions

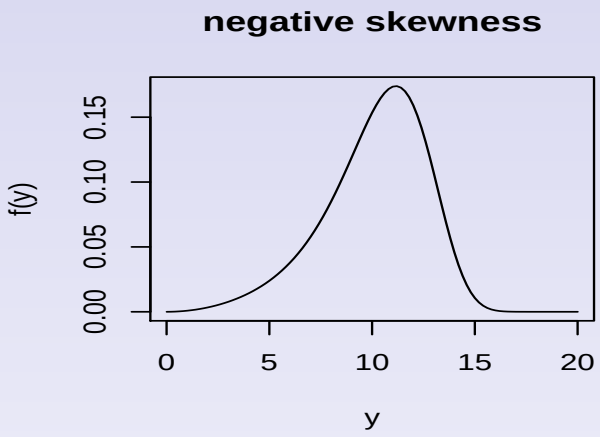


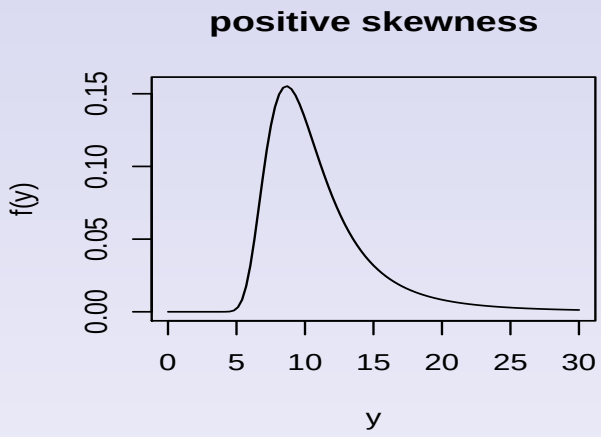


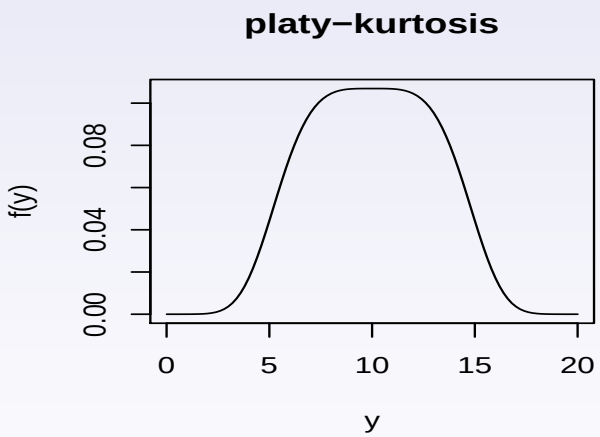
gamlss

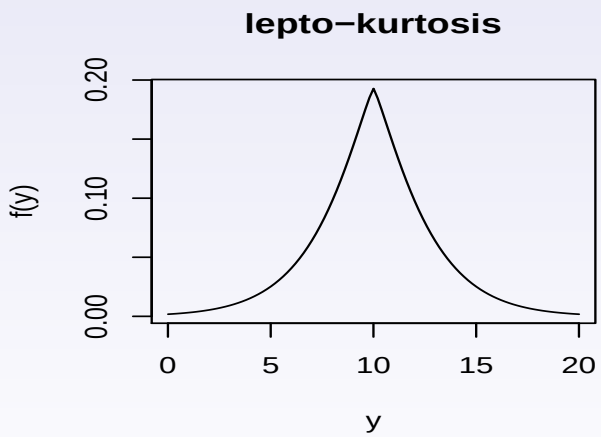
Continuous distributions: different shapes

Go back Distributions





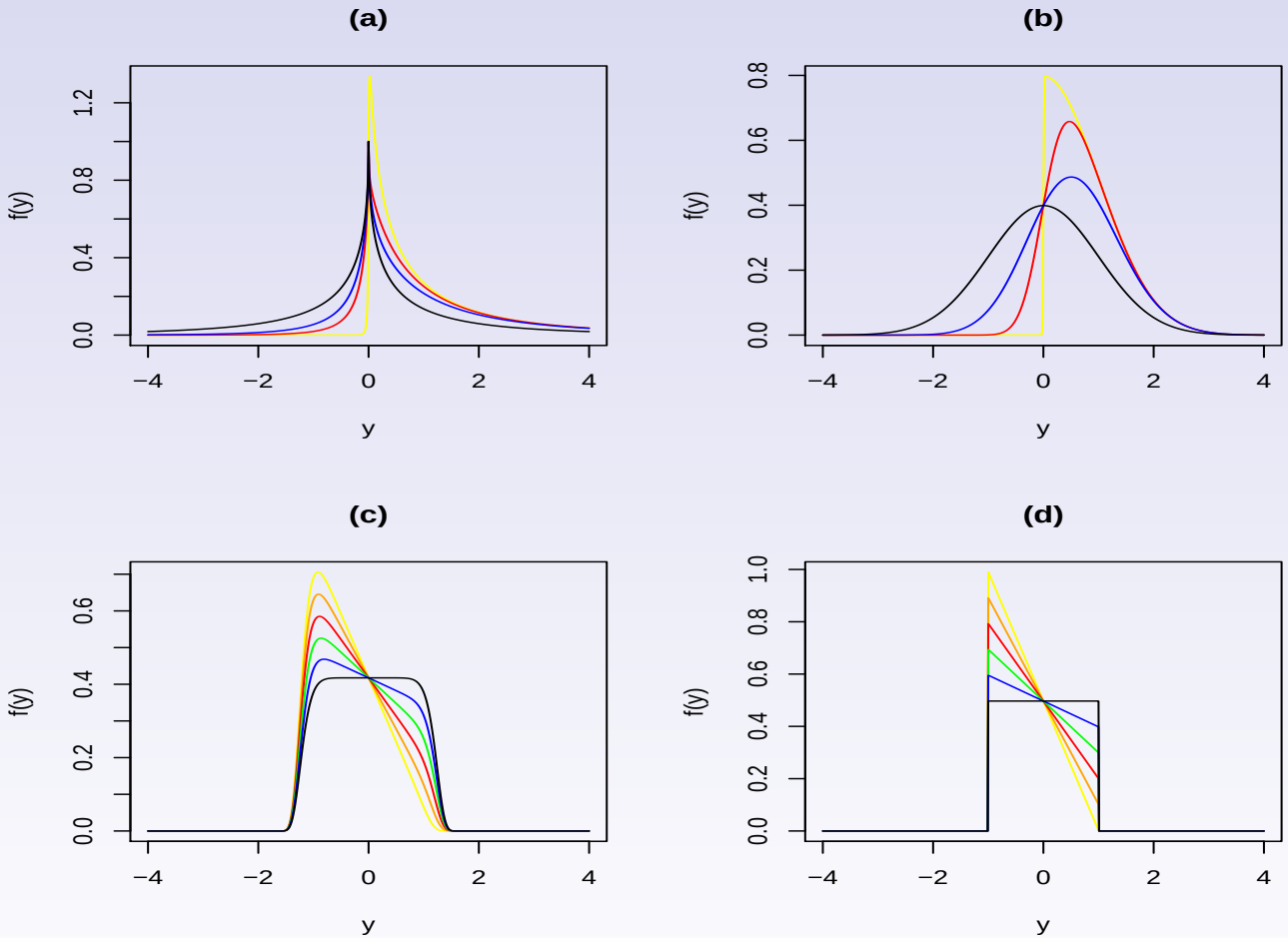




gamlss

Continuous distributions: different types

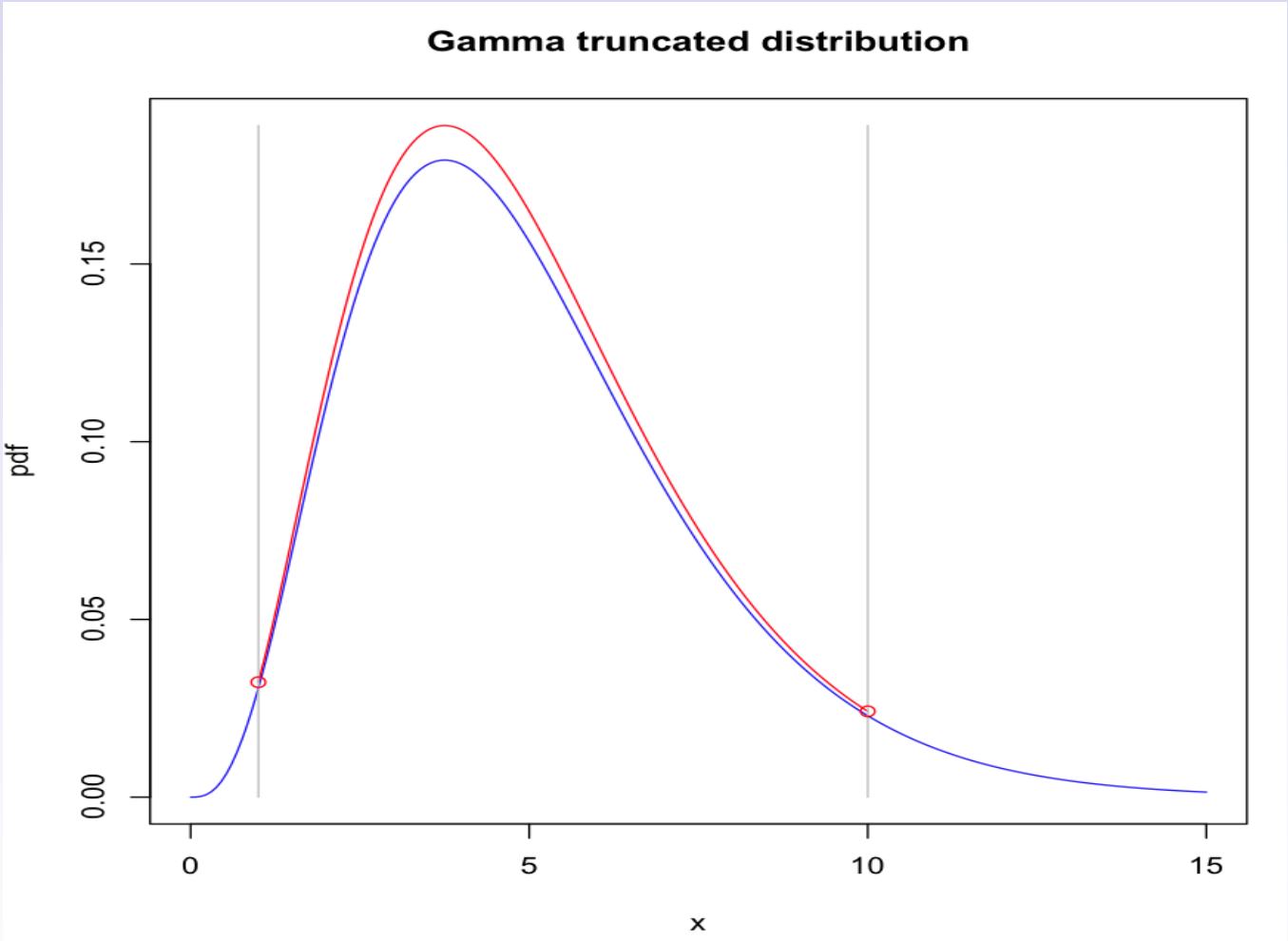
Go back Distributions



gamlss

Continuous distributions: different types

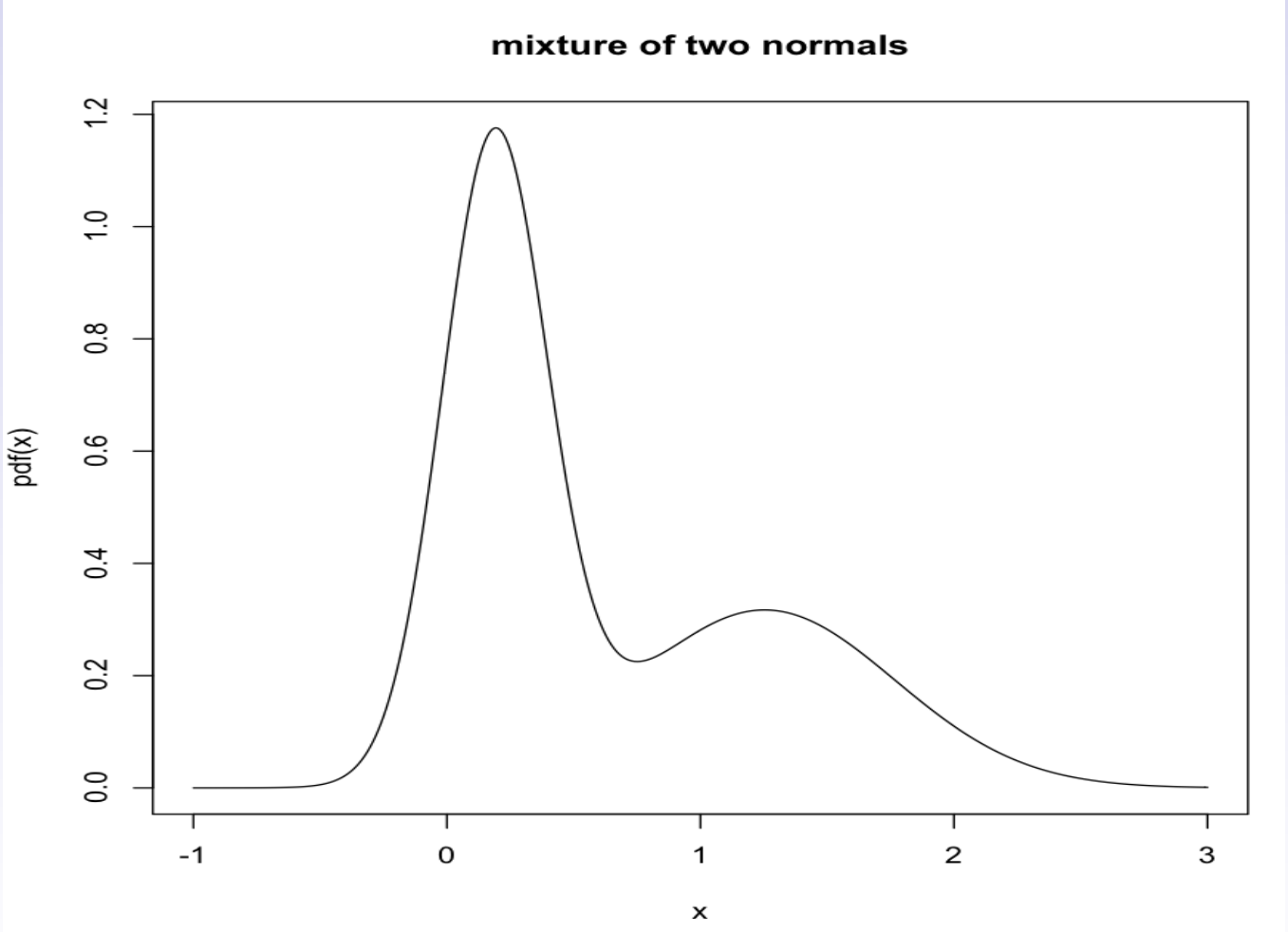
Go back Distributions



gamlss

Continuous distributions: different types

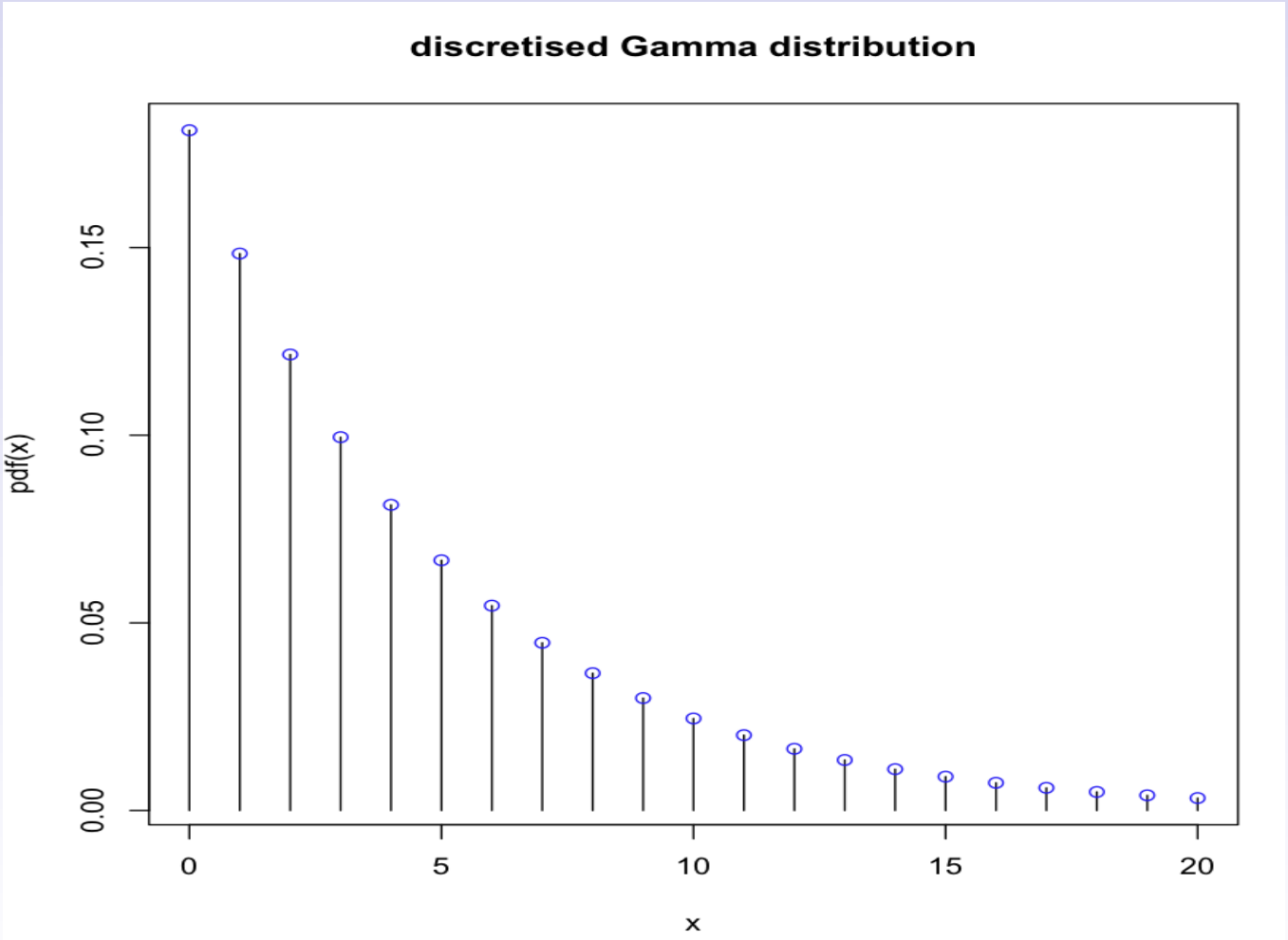
Go back Distributions



gamlss

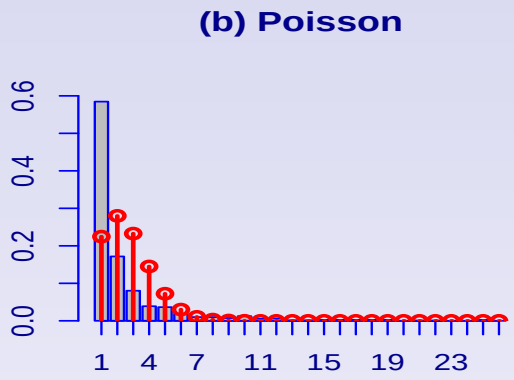
Continuous distributions: different types

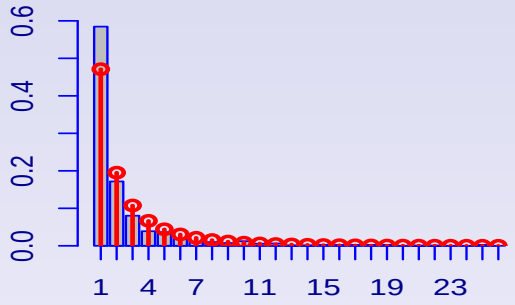
Go back Distributions

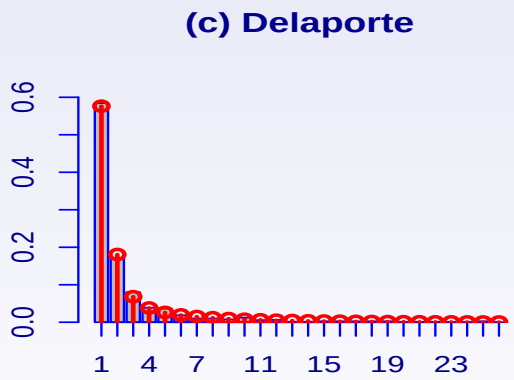


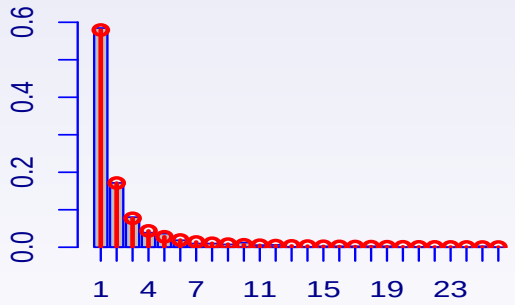
gamlss

The stylometric data, Go back to distributions









gamlss