Gamma regression models with the Gammareg R package[1]

Regresión Gamma con el paquete Gammareg en R

Marta Corrales Bossio[a]
martac@unal.edu.co

Edilberto Cepeda Cuervo[b]
ecepedac@unal.edu.co

Resumen

The class of gamma regression models is based on the assumption that the dependent variable is gamma distributed and that its mean is related to a set of regressors through a linear predictor with unknown coefficients and a link function. This link can be the identity, the inverse or the logarithm function. The model also includes a shape parameter, which may be constant or dependent on a set of regressors through a link function, as the logarithm function. In this paper we describe the Gammareg R-package, which provides the class of gamma regressions in the R system for their statistical computing. The underlying theory is briefly presented and the library implementation illustrated from simulation studies.

Palabras clave: Gamma regression, mean regression structures, shape regression structures, Fisher Scoring algorithm, R-package.

Abstract

En este artículo se presenta el paquete Gammareg, el cual utiliza el método Fisher-Scoring para ajustar modelos de regresión gamma, donde tanto la media como el parámetro de forma poseen estructuras de regresión, y el cual fue desarrollado en el *software* R. Después de realizar una breve presentación de la teoría subyacente, se presenta el uso de la librería por medio de estudios de simulación y aplicaciones.

Keywords: Algoritmo Fisher-Scoring, estructura de regresión en forma, estructura de regresión en media, regresión Gamma, software R.

---

[a]Docente. Universidad Sergio Arboleda. Colombia.
[b]Profesor titular. Departamento de Estadística. Universidad Nacional de Colombia.

1. Introduction

The gamma distribution can be used with great flexibility in the analysis of positive random variables. Thus, gamma regression models are applied in a wide range of empirical applications, such as in the process of rate setting in the frame-work of heterogeneous insurance portfolios, which is the most important function of insurers (Krishnamoorthy, 2006), and in hospital admissions for rare diseases where time series are very sparse due to infrequency of events (Winklemann, 2008).

In this paper we consider gamma regression models in which both mean and shape parameters are allowed to depend on covariates. These models, in which these two parameters follow regression structures, are proposed in Cepeda-Cuervo (2001) under both classic and Bayesian approaches. We introduce the use of the Gammareg R-package, an R-code developed to fit gamma regression models, applying an alternative iterative maximum likelihood method based on the Fisher scoring algorithm.

After the introduction, this paper includes six sections. In Section 2 the mean-shape re-parameterizations of the gamma distribution are presented. In Section 3 the gamma regression models, where both mean and shape parameters follow regression structures, are presented. In Section 4 a Fisher scoring method to fit gamma regression models is summarized. Section 5 presents the Gammareg R-package. Section 6 contains two applications based on simulated data. Finally, in Section 7 our main conclusions are presented.

2. Gamma distribution

A random variable Y follows a gamma distribution if its probability density function is given by

$$f(y|\alpha, \lambda) = \frac{\lambda^\alpha y^{\alpha-1} e^{-\lambda y}}{\Gamma(\alpha)} \boldsymbol{I}_{(0,\infty)}(y), \tag{1}$$

where $\alpha, \lambda > 0$, $\Gamma(.)$ denotes the gamma function, and $\boldsymbol{I}(.)$ is the indicator function. Under this parameterization, the mean and variance of Y are given by $\mathrm{E}(Y) = \alpha/\lambda$ and $\mathrm{Var}(Y) = \alpha/\lambda^2 = \mu^2/\alpha$. Given that $\lambda = \alpha/\mu$, the gamma distribution function (1) can be re-parameterized as a function of the mean (μ) and shape (α) parameters, as proposed in Cepeda-Cuervo (2001) and Cepeda-Cuervo & Gammerman (2005), where the gamma distribution function is written as in equation (2).

$$f(y|\mu, \alpha) = \frac{1}{y\Gamma(\alpha)} \left(\frac{\alpha y}{\mu}\right)^\alpha e^{-\alpha y/\mu} \boldsymbol{I}_{(0,\infty)}(y) \tag{2}$$

where $\mu, \alpha > 0$, $\Gamma(.)$ denotes the gamma function, and $\boldsymbol{I}(.)$ is the indicator function. The notation $Y \sim G(\mu, \alpha)$ is used to denote that Y follows a gamma distribution with $\mathrm{E}(Y) = \mu$ and α as a shape parameter.

Using this re-parameterization of the gamma distribution, the joint mean and shape gamma regression models proposed in Cepeda-Cuervo (2001), is presented in the next section.

3. Gamma regression models

Let $Y_i \sim G(\mu_i, \alpha)$, $i = 1, \ldots, n$, be a random sample of size n. In gamma regression models with constant shape parameter, the mean regression structure is defined by

$$g(\mu_i) = \eta_i = \boldsymbol{x}_i'\boldsymbol{\beta},$$

where g is the link function, $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_p)'$ is the vector of mean regression parameters, $\boldsymbol{x}_i$ is the i-th vector value of the explanatory variables, and η_i is a linear predictor. Here, $g(.) : (0, \infty) \mapsto \Re$ is a real value function, strictly monotonic and twice differentiable (McCullagh and Nelder, 1989). Some usual mean link functions in the gamma regression are: the logarithm function, $g(\mu) = log(\mu)$; the identity function, $g(\mu) = \mu$, and the inverse function $g(\mu) = 1/\mu$. In generalized linear models, the canonical link for the mean is the inverse function.

An extension of the gamma regression models is proposed in Cepeda-Cuervo (2001). In this proposal, the shape parameter is not constant through the observations and is modeled following a regression structure. That is, the mean and shape parameters follow the regression structures given by equations (3) and (4):

$$g(\mu_i) \quad = \quad \eta_{1i} = \boldsymbol{x}_i'\boldsymbol{\beta} \tag{3}$$

$$h(\alpha_i) \quad = \quad \eta_{2i} = \boldsymbol{z}_i'\boldsymbol{\gamma} \tag{4}$$

where g and h are appropriate real link functions; $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_p)'$ and $\boldsymbol{\gamma} = (\gamma_0, \ldots, \gamma_k)'$ are, respectively, the mean and the shape regression parameter vectors; $\boldsymbol{x_i}$ and $\boldsymbol{z_i}$ are, respectively, the mean and the shape explanatory variables for the i-th observation and, η_{1i} and η_{2i} are the linear predictors. A usual link function in the shape regression structure is the logarithm function.

4. The method to fit gamma regression models

Cepeda-Cuervo (2001) propose a classic approach to fit gamma regression models where both mean and shape parameters follow regression structures, using the Fisher scoring algorithm. In that work, he shows that with the gamma reparameterization given by (2) the likelihood function can be written in the form:

$$L(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \prod_{i=1}^{n} \frac{1}{\Gamma(\alpha_i)} \left(\frac{\alpha_i}{\mu_i}\right)^{\alpha_i} y_i^{\alpha_i - 1} \exp\left(-\frac{\alpha_i}{\mu_i} y_i\right) \tag{5}$$

And the log-likelihood function by:

$$l(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \sum_{i=1}^{n} \left\{ -\log[\Gamma(\alpha_i)] + \alpha_i \log\left(\frac{\alpha_i y_i}{\mu_i}\right) - \log(y_i) - \left(\frac{\alpha_i}{\mu_i}\right) y_i \right\} \qquad (6)$$

Thus, assuming the regression structures defined by $\mu_i = \boldsymbol{x}_i' \boldsymbol{\beta}$ and $\alpha_i = \exp(\boldsymbol{z}_i' \boldsymbol{\gamma})$, the score statistics are given by:

$$\frac{\partial l}{\partial \beta_j} = \sum_{i=1}^{n} -\frac{\alpha_i}{\mu_i} \left(1 - \frac{y_i}{\mu_i}\right) x_{ij}, \quad j = 1, \ldots p$$

$$\frac{\partial l}{\partial \gamma_k} = \sum_{i=1}^{n} -\alpha_i \left[\frac{d}{d\alpha_i} \log \Gamma(\alpha_i) - \log\left(\frac{\alpha_i y_i}{\mu_i}\right) - 1 + \frac{y_i}{\mu_i}\right] z_{ik}, \quad k = 1, \ldots, r$$

and the Hessian matrix is determined by:

$$\frac{\partial^2 l}{\partial \beta_k \partial \beta_j} = \sum_{i=1}^{n} \frac{\alpha_i}{\mu_i^2} \left(1 - \frac{2y_i}{\mu_i}\right) x_{ij} x_{ik}, \quad j, k = 1, \ldots p$$

$$\frac{\partial^2 l}{\partial \gamma_k \partial \beta_j} = \sum_{i=1}^{n} \frac{\alpha_i}{\mu_i} \left[-1 + \frac{y_i}{\mu_i}\right] x_{ij} z_{ik}, \quad j = 1, \ldots, p, \ k = 1, \ldots, r$$

$$\frac{\partial^2 l}{\partial \gamma_k \partial \gamma_j} = \sum_{i=1}^{n} -\alpha_i \left[\frac{d}{d\alpha_i} \log \Gamma(\alpha_i) - \log\left(\frac{\alpha_i y_i}{\mu_i}\right) - 1 + \frac{y_i}{\mu_i}\right] z_{ij} z_{ik}$$
$$- \sum_{i=1}^{n} \alpha_i \left[\alpha_i \frac{d^2}{d\alpha_i^2} \Gamma(\alpha_i) - 1\right] z_{ij} z_{ik}, \quad j, k = 1, \ldots, r$$

Thus, the Fisher information matrix is given by:

$$-E\left(\frac{\partial^2 l}{\partial \beta_k \beta_j}\right) = \sum_{i=1}^{n} \frac{\alpha_i}{\mu_i^2} x_{ij} x_{ik}, \quad j, k = 1, \cdots, p$$

$$-E\left(\frac{\partial^2 l}{\partial \gamma_k \beta_j}\right) = 0, \quad j = 1, \cdots, p, \ k = 1, \cdots, r$$

$$-E\left(\frac{\partial^2 l}{\partial \beta_k \beta_j}\right) = \sum_{i=1}^{n} \alpha_i^2 \left[\frac{d^2}{d\alpha_i^2} \log \Gamma(\alpha_i) - \frac{1}{\alpha_i}\right] z_{ij} z_{ik}, \quad j, k = 1, \cdots, r$$

It can be noted that the Fisher information matrix is a block diagonal matrix, where one of the blocks corresponds to the mean regression parameters $\boldsymbol{\beta}$ and the other to the shape regression parameter $\boldsymbol{\gamma}$. Thus, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are orthogonal

(Cox & Reid 1987) and their maximum likelihood estimators are asymptotically independent.

Finally, taking in account the structure of the Fisher information matrix, Cepeda-Cuervo (2001) shows that the Fisher scoring information equation can be written as following equation system:

$$\boldsymbol{\beta}^{(k+1)} \;=\; (X'W_1^{(k)}X)^{-1}X'W_1^{(k)}Y \tag{7}$$

$$\boldsymbol{\gamma}^{(k+1)} \;=\; (Z'W_2^{(k)}Z)^{-1}X'W_2^{(k)}\tilde{Y} \tag{8}$$

where $\boldsymbol{W}_1^{(k)}$ is a diagonal matrix with diagonal entries $w_{ii}^{(k)} = (\mu_i^2/\alpha_i)$, and

$$\tilde{y}_i = \eta_{2i} - \frac{1}{\alpha_i}\Big[\frac{\partial^2}{\partial\alpha^2}\mathrm{log}\Gamma(\alpha_i) - \frac{1}{\alpha_i}\Big]^{-1}\Big[\frac{\partial}{\partial\alpha_i}\mathrm{log}\Gamma(\alpha_i) - \mathrm{log}\Big(\frac{\alpha_i y_i}{\mu_i}\Big) - 1 + \frac{y_i}{\mu_i}\Big].$$

$\boldsymbol{W}_2^{(k)}$ is a diagonal matrix with diagonal entries $w_{ii}^{(k)} = 1/d_i$, with

$$d_i = \alpha_i^{-2}\Big[\frac{d^2}{d\alpha_i^2}\log\Gamma(\alpha_i) - \frac{1}{\alpha_i}\Big]^{-1}.$$

Thus, based on the structure of the Fisher information matrix, Cepeda-Cuervo (2001) proposes an iterative algorithm to obtain the maximum likelihood estimates of the regression parameters, as follows:

1. Start an iteration counter at $k = 0$.

2. Give initial values for the parameters β and γ.

3. Obtain $\boldsymbol{\beta}^{(k+1)}$ from equation (7), given the current values of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$.

4. Obtain $\boldsymbol{\gamma}^{(k+1)}$ from equation (8), given the current values of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$.

5. Set the counter iteration $k = k + 1$

6. Go to 3 until convergence.

For other link functions, like $g(.) = log(.)$ and $h(.) = log(.)$, the Fisher information matrix is block diagonal and a similar alternate iterated algorithm is implemented.

5. Implementation in R: Gammareg package

The gamma regression models as proposed in McCullagh and Nelder (1989) can be fitted using the R-function glm {stats} of the free statistical software R (Development Core Team, 2009). The Gammareg R-package is the computational implementation of the maximum likelihood method presented in Section 4, to fit the gamma regression models where both mean and shape parameters follow regression structures. The main model-fitting function in **Gammareg** is Gammareg(),

which allows the user to estimate the mean and shape regression parameters in a gamma regression model from the classic perspective. The general formula for this function is:

$$\textbf{\textit{Gammareg(formula1, formula2, meanlink)}}$$

where

1. **formula1** is an object of class formula, that describes the variable of interest Y and the regressors X of the mean structure.

2. **formula2** is an object of class formula, that describes the regressors Z for the shape structure, with log-link function.

3. **meanlink** is the link function for the mean.

The default mean link function is the logarithm (log), but the identity function (ide) is also admissible.

The fitted-model object of the Gammareg class is a list similar to **glm** objects. It provides to the user the regression parameter estimates ($\widehat{\boldsymbol{\beta}}$ and $\widehat{\boldsymbol{\gamma}}$) and their standard deviations. It also provides the estimated covariance matrix for $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, the AIC value, the number of iterations to convergence and the convergence value obtained.

The **Gammareg** R-package has five other functions which allow the user, among other things, to obtain summaries for gamma regression models. These functions are described in Table 1.

Tabla 1: *Gammareg functions*

Function	Description
Gammareg	Estimates the mean and shape regression parameters.
gammahetero1()	Performs the classic gamma regression using link log for the mean and link log for the shape.
gammahetero2()	Performs the classic gamma regression using identity link for the mean and link log for the shape
summary.Gammareg()	Is the standard regression output (coefficient estimates, standard errors, criteria); returns an object of class summary. Gammareg containing the relevant summary statistics (which has a print() method)
gammaresiduals	Calculates the Deviance, Pearson, Standarized Pearson residuals and a special gamma residual (r^*) based in the properties of the biparametric exponential family distribution.
print.Gammareg	Prints the estimated coefficients and their confidence intervals.
print.summary.Gammareg	Prints the summary of a classic gamma regression.

6. Gammareg in practice

To illustrate the use of the **Gammareg** R-package, here we present the results obtained by fitting two gamma regression models to a simulated dataset. In the

first study the identity mean link function is assumed and in the second the mean logarithm function is assumed.

6.1. First simulation

In a first simulation, a gamma regression model with mean and shape structures given by equations (9) and (10), respectively, is considered:

$$\mu_i = \boldsymbol{x_i}'\boldsymbol{\beta} \qquad (9)$$

$$\log(\alpha_i) = \boldsymbol{z_i'}\boldsymbol{\gamma} \qquad (10)$$

In this simulation, four covariates X_i, $i = 1, \ldots, 4$ are considered. $n = 500$ values were simulated for each covariate. X_1 is assumed to be a constant variable, with $x_{1i} = 1$ for all i, in order to consider an intercept in the regression structures. Values of the covariate X_2 were generated from a uniform distribution in the interval $(0, 30)$; values of X_3 were generated from a uniform distribution in the interval $(0, 15)$; and values of X_4 were generated from a uniform distribution in the interval $(10, 20)$. The values Y_i were generated from a gamma distribution with mean and shape parameters given by $\mu_i = 15 + 2x_{2i} + 3x_{3i}$ and $\alpha_i = exp(0.2 + 0.1x_{2i} + 0.3x_4i)$, as follows:

```
>library(Gammareg)
>X1 <- rep(1, 500)
>X2 <- runif(500, 0, 30)
>X3 <- runif(500, 0, 15)
>X4 <- runif(500, 10, 20)
>mui <- 15 + 2*X2 + 3*X3
>alphai <- exp(0.2 + 0.1*X2 + 0.3*X4)
>Y <- rgamma(500, shape=alphai, scale=mui/alphai)
>X <- cbind(X1, X2, X3)
>Z <- cbind(X1, X2, X4)
>formula.mean = Y ~ X2 + X3
>formula.shape = ~ X2 + X4
>a=Gammareg(formula.mean, formula.shape, meanlink="ide")
>summary(a)
```

The results obtained to apply the Gammareg R-package were:

```
###################################################################
###                Classic Gamma Regression                 ###
###################################################################

Call:
Gammareg(formula1 = formula.mean, formula2 = formula.shape,
    meanlink = "ide")

                  Estimate  L.Intv U.Intv
beta.(Intercept)   15.2231 14.6931 15.753
beta.X2             2.0382  2.0131  2.063
beta.X3             2.8965  2.8483  2.945
gamma.(Intercept)   0.1877  0.1870  0.189
gamma.X2            0.1031  0.1031  0.103
gamma.X4            0.2934  0.2933  0.293

Covariance Matrix for Beta:
            (Intercept)            X2            X3
(Intercept)  0.072747993 -2.615207e-03 -2.819056e-03
X2          -0.002615207  1.637345e-04 -3.582766e-05
X3          -0.002819056 -3.582766e-05  6.024330e-04

Covariance Matrix for Gamma:
            (Intercept)            X2            X4
(Intercept)  1.485298e-07 -1.243918e-09 -6.456682e-09
X2          -1.243918e-09  4.430374e-11  7.859803e-12
X4          -6.456682e-09  7.859803e-12  3.468175e-10

AIC:
[1] 6606828

Iteration:
[1] 13

Convergence:
[1] 4.660965e-06
```

The results show that all the parameter estimates are close to the true parameter values of the model. In all cases, the standard deviations are small and the 95 % confidence interval contains the true parameter value.

6.2. Second simulation

In a second simulation, a gamma regression model with mean and shape structures given by equations (11) and (12), respectively, is considered:

$$\log(\mu_i) = \boldsymbol{x}_i'\beta \tag{11}$$
$$\log(\alpha_i) = \boldsymbol{z}_i'\gamma \tag{12}$$

In this case, $n = 500$ values of the explanatory variables X_i, $i = 1, \ldots, n$, were generated as in the first simulation, but the values of the variable of interest Y were generated from a gamma distribution with $\mu_i = -5 + 2x_{2i} + 3x_{3i}$ and $\alpha_i = \exp(0.2 + 0.1x_{2i} + 0.3x_4 i)$ as follows:

```
>library(Gammareg)
X1 <- rep(1, 500)
X2 <- runif(500, 0, 30)
X3 <- runif(500, 0, 15)
X4 <- runif(500, 10, 20)
mui <- exp(-5 + 0.2*X2 -0.03*X3)
alphai <- exp(0.2 + 0.1*X2 + 0.3*X4)
Y <- rgamma(500, shape=alphai, scale= mui/alphai)
X <- cbind(X1, X2, X3)
Z <- cbind(X1, X2, X4)
formula.mean = Y ~ X2 + X3
formula.shape = ~ X2 + X4
a=Gammareg(formula.mean, formula.shape, meanlink="log")
summary(a)
```

The results obtained from the application of the Gammareg R-package were:

```
####################################################################
###                    Classic Gamma Regression               ###
####################################################################

Call:
Gammareg(formula1 = formula.mean, formula2 = formula.shape,
meanlink = "log")

                  Estimate   L.Intv U.Intv
beta.(Intercept)  -4.98768 -4.99730 -4.978
beta.X2            0.19953  0.19915  0.200
beta.X3           -0.03021 -0.03083 -0.030
gamma.(Intercept)  0.33678  0.33610  0.337
gamma.X2           0.08687  0.08686  0.087
```

```
gamma.X4              0.30849  0.30846  0.309

 Covariance Matrix for Beta:
              (Intercept)              X2              X3
(Intercept)  2.397679e-05 -7.517510e-07 -8.062176e-07
X2           -7.517510e-07  3.746544e-08 -1.793592e-09
X3           -8.062176e-07 -1.793592e-09  1.016449e-07

 Covariance Matrix for Gamma:
              (Intercept)              X2              X4
(Intercept)  1.188862e-07 -8.320284e-10 -5.352878e-09
X2           -8.320284e-10  2.967992e-11  7.473412e-12
X4           -5.352878e-09  7.473412e-12  2.810451e-10

 AIC:
[1] -3469938

 Iteration:
[1] 20

 Convergence:
[1] 1.085736e-08
```

In this case, these results also show that all the parameter estimates obtained using the Gammareg R-package are close to the true parameter values of the model, but in all cases the 95 % confidence interval contains the true parameter value, except for γ_0. The standard deviations are small for all parameters.

6.3. Application: Duration of the embryonic stage of fruit flies

In this application, we analyze a dataset collected by (Powsner 1935), as part of an experiment to determine accurately the effect of temperature on the duration of the developmental stages of the fruit fly (Drosophila melanogaster). In his experiment, Powsner studied four stages: the embryonic, egg-larval, larval and pupal stages.

This dataset was analyzed by McCullagh and Nelder (1989), assuming a gamma regression model where observed duration is the response variable weighted by batch size, in the first stage (embryonic). Although, they considered different rational mean functions of temperature, all of them with constant shape parameter, we assume a gamma regression model where both mean and shape parameters follow regression structures given by:

$$\log(\mu_i) = \exp\left(\beta_0 + \beta_1 T_i + \frac{\beta_2}{T_i}\right)$$

$$\log(\alpha_i) = \exp\left(\lambda_0 + \lambda_1 T_i + \frac{\lambda_2}{T_i}\right)$$

with $\delta = 58.64$ as in McCullagh and Nelder (1989).

The results obtained from the application of the Gammareg R-package were:

```
##################################################################
###                 Classic Gamma Regression              ###
##################################################################

 Call:
Gammareg(formula1 = formula.mean, formula2 = formula.shape,
meanlink = "log")

                   Estimate    L.Intv   U.Intv
beta.(Intercept)    3.22001     0.5653   5.087
beta.X2            -0.2796     -0.6427  -0.016
beta.X3          -219.7723   -659.6103 -96.066
gamma.(Intercept)   1.1511     -2.2147  -0.517
gamma.X2           -0.0957     -0.2627  -0.001
gamma.X3          -70.4925   -299.1699 -58.185

 Covariance Matrix for Beta:
            (Intercept)          X2           X3
(Intercept)    2.342089  0.25226402      0.46954
X2             0.252264  0.02238432     26.76824
X3             0.469537 26.76824173   3588.39007

 Covariance Matrix for Gamma:
             (Intercept)          X2           X3
(Intercept)   1.27767636 0.094231012    119.928055
X2            0.09423101 0.007986262      9.681734
X3          119.92805461 9.681733520   1137.071768

 AIC:
[1] -612.2067

 Iteration:
[1] 9

 Convergence:
[1] 7.290849e-06
```

We obtained a good fit of the assumed model, as noted in Figure 1, where the dots represent the observed values and the segmented line the fitted mean.
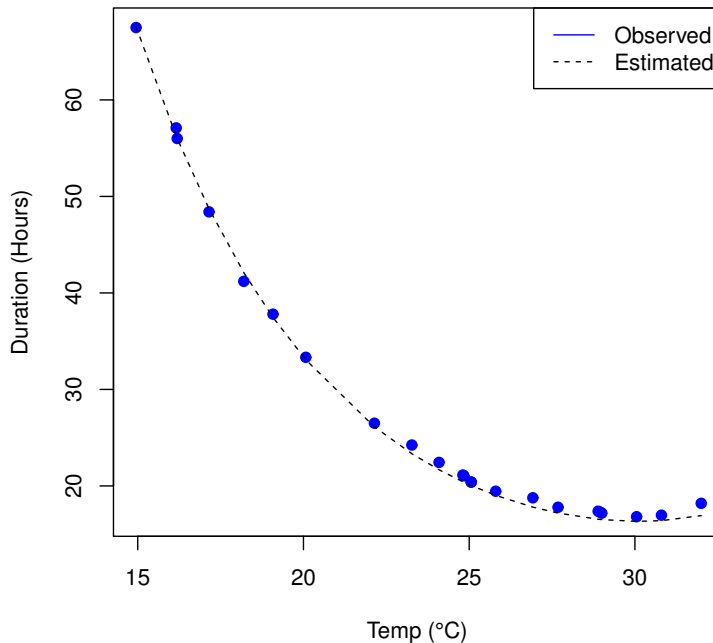

Figura 1: *Fitted curve (—) and observed duration ·. Source: Own elaboration.*

7. Conclusions and extensions

This paper introduces the Gammareg R-package, which can be used to fit gamma regression models by applying the classic method proposed in Cepeda-Cuervo (2001). Two simulated studies are used to illustrate the use of the different functions of this package. In all of cases, there is a fast convergence to the maximum likelihood estimation of the regression parameters, showing the efficiency of the alternate block-iterative algorithm.

There are many possibilities for future works and practical issues. One is the use of alternative link functions, like the inverse link, which fit some database better. Another is to develop an R-package to fit the Bayesian gamma regression models proposed in Cepeda-Cuervo (2001), where both mean and variance follow regression structures. These studies are in development.

Referencias

Bateson, T. F. (2009), 'Gamma regression of interevent waiting times versus poisson regression of daily event counts: Inside the epidemiologist's toolbox selecting the best modeling tools for the job', *Epidemiology* **20**(2), 202–204.

Cepeda-Cuervo, E. (2001), 'Modelagem de variabilidade em modelos lineares generalizados', *Unpublished Ph.D.thesis, Mathematics Institute, Universidade Federal Rio de Janeiro* .

Cepeda, E. & Gammerman, D. (2005), 'Bayesian methodology for modeling parameters in the two parameters exponential family', *ESTADISTICA* **57**(168), 93–105.

Chib, S. & Greenberg, E. (1995), 'Understanding the metropolis-hastings algorithm', *The American Statistician* **49**(4), 327–335.

Gamerman, D. & Lopes, H. F. (2006), *Markov chain Monte Carlo: Stochastic simulation for Bayesian inference*, CRC Press, address=New York,.

Krishnamoorthy, K. (2006), *Handbook of Statistical Distributions with Applications*, Chapman & Hall/CRC, Florida.

McCullagh, J. & Nelder, J. (1989), *Generalized Linear Models. Second Edition*, Chapman and Hall, London.

Team., D. C. (n.d.), *A language and environment for statistical computing. R Foundation for Statistical Computing.*

Winklemann, R. (2008), *Econometric analysis of count data*, Springer-Verlag, Berlin, Germany.