

# GLM Benchmarks

```
rm(list=ls())
#install.packages("glm2")
#install.packages("alpaca")
#install.packages("FENmlm")
library(alpaca)
library(glm2)
library(FENmlm)
setwd("C:/Git/ppml_hdfe_demo/guides")
```

## Base R: glm

```
rm(list=ls())
data <- read.csv(file="csv/example1.csv", header=TRUE, sep=",")
formula <- y ~ x1 + x2 + x3 + x4
mod <- glm(formula, data, family=poisson())
summary(mod)

##
## Call:
## glm(formula = formula, family = poisson(), data = data)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.97355  -0.75131  -0.16879   0.07357   2.70708
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   0.59095    0.30291   1.951  0.0511 .
## x1          -17.78017   3467.85856  -0.005  0.9959
## x2           17.32952   3467.85857   0.005  0.9960
## x3           -0.47085    0.23117  -2.037  0.0417 *
## x4           -0.03779    0.04375  -0.864  0.3878
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##      Null deviance: 31.912  on 11  degrees of freedom
## Residual deviance: 15.956  on 7  degrees of freedom
## AIC: 46.991
##
## Number of Fisher Scoring iterations: 15

rm(list=ls())
data <- read.csv(file="csv/example2.csv", header=TRUE, sep=",")
formula <- y ~ x1 + x2 + x3 + x4
mod <- glm(formula, data, family=poisson())

## Warning: glm.fit: algorithm did not converge
```

```
## Warning: glm.fit: fitted rates numerically 0 occurred
summary(mod)

##
## Call:
## glm(formula = formula, family = poisson(), data = data)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -3.109e-03 -2.000e-08 -2.000e-08 -2.000e-08  2.000e-08
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -367.83    7448.51  -0.049   0.961
## x1             512.42   10345.15   0.050   0.960
## x2          -1644.86   33284.96  -0.049   0.961
## x3           -105.85    2138.00  -0.050   0.961
## x4              20.56     413.81   0.050   0.960
##
## (Dispersion parameter for poisson family taken to be 1)
##
##      Null deviance: 4.6719e+01  on 11  degrees of freedom
## Residual deviance: 9.6716e-06  on  7  degrees of freedom
## AIC: 19.933
##
## Number of Fisher Scoring iterations: 25
```

## glm2

```
rm(list=ls())
data <- read.csv(file="csv/example1.csv", header=TRUE, sep=",")
formula <- y ~ x1 + x2 + x3 + x4
mod <- glm2(formula, data, family=poisson())
summary(mod)

##
## Call:
## glm2(formula = formula, family = poisson(), data = data)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.97355  -0.75131  -0.16879   0.07357   2.70708
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   0.59095    0.30291   1.951  0.0511 .
## x1          -17.78017   3467.85856  -0.005  0.9959
## x2           17.32952   3467.85857   0.005  0.9960
## x3           -0.47085    0.23117  -2.037  0.0417 *
## x4           -0.03779    0.04375  -0.864  0.3878
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## (Dispersion parameter for poisson family taken to be 1)
##
## Null deviance: 31.912 on 11 degrees of freedom
## Residual deviance: 15.956 on 7 degrees of freedom
## AIC: 46.991
##
## Number of Fisher Scoring iterations: 15
rm(list=ls())
data <- read.csv(file="csv/example2.csv", header=TRUE, sep=",")
formula <- y ~ x1 + x2 + x3 + x4
mod <- glm2(formula, data, family=poisson())

## Warning: glm.fit2: algorithm did not converge. Try increasing the maximum
## iterations

## Warning: glm.fit2: fitted rates numerically 0 occurred
summary(mod)

##
## Call:
## glm2(formula = formula, family = poisson(), data = data)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -3.109e-03 -2.000e-08 -2.000e-08 -2.000e-08  2.000e-08
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -367.83     7448.51  -0.049   0.961
## x1             512.42    10345.15   0.050   0.960
## x2          -1644.86    33284.96  -0.049   0.961
## x3           -105.85     2138.00  -0.050   0.961
## x4              20.56      413.81   0.050   0.960
##
## (Dispersion parameter for poisson family taken to be 1)
##
## Null deviance: 4.6719e+01 on 11 degrees of freedom
## Residual deviance: 9.6716e-06 on 7 degrees of freedom
## AIC: 19.933
##
## Number of Fisher Scoring iterations: 25
```

## alpaca

```
rm(list=ls())
data <- read.csv(file="csv/fe1.csv", header=TRUE, sep=",")
formula <- y ~ x1 + x2 | i + j
mod <- feglm(formula, data, family=poisson())
summary(mod)

## poisson
```

```
##
## y ~ x1 + x2 | i + j
##
## l= [4, 4], n= 18, deviance= 13.3644
##
## Structural parameter(s):
##
##      Estimate Std. error z value Pr(> |z|)
## x1   -0.4845     1.2439  -0.390   0.697
## x2  -18.9906   3934.0725  -0.005   0.996

rm(list=ls())
data <- read.csv(file="csv/fe2.csv", header=TRUE, sep=",")
formula <- y ~ x1 | i + j
try(mod <- feglm(formula, data, family=poisson()))

## Error in feglm(formula, data, family = poisson()) :
##   Backtracking (step-halving) failed.

#summary(mod)
```

## FENmlm

```
rm(list=ls())
data <- read.csv(file="csv/fe1.csv", header=TRUE, sep=",")
formula <- y ~ x1 + x2 | i + j
mod <- FENmlm::femlm(formula, data, family="poisson")

## Warning: [Getting cluster coefficients] iteration limit reached (10000).
## Warning: [Getting cluster coefficients] iteration limit reached (10000).
## Warning: [Getting cluster coefficients] iteration limit reached (10000).
## Warning: [Getting cluster coefficients] iteration limit reached (10000).
## Warning: [Getting cluster coefficients] iteration limit reached (10000).
## Warning: [Getting cluster coefficients] iteration limit reached (10000).
## Warning in FENmlm::femlm(formula, data, family = "poisson"): [femlm]: The
## optimization algorithm did not converge, the results are not reliable. Use
## function diagnostic() to see what's wrong.
## Warning: [Getting cluster coefficients] iteration limit reached (10000).

summary(mod)

## Warning in print.femlm(x): The optimization algorithm did not converge, the
## results are not reliable. Use function diagnostic() to see what's wrong.
## ML estimation, family = Poisson, Dep. Var.: y
## Observations: 18
## Cluster sizes: i: 4, j: 4
## Standard-errors type: Standard
```

```
##      Estimate Std. Error   z value Pr(>|z|)
## x1  -0.515619     1.2564 -0.410396 0.681515
## x2 -15.094000    1467.4000 -0.010286 0.991793
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##      BIC: -12.33      Pseudo-R2: -0.33391
## Log-likelihood:  76.70    Squared Cor.: 0.26944
## # Evaluations: 16 -- false convergence (8)

rm(list=ls())
data <- read.csv(file="csv/fe2.csv", header=TRUE, sep=",")
formula <- y ~ x1 | i + j
try(mod <- FENmlm::femlm(formula, data, family="poisson"))

## Warning in FENmlm::femlm(formula, data, family = "poisson"): 1/2 clusters
## (5 observations) removed because of only zero outcomes.

## Error in if (ll == (-Inf)) return(1e+308) :
##   missing value where TRUE/FALSE needed

#summary(mod)
#diagnostic(mod)
```