

# Modèles Linéaires Appliqués / Régression

## Agenda

Arthur Charpentier

UQAM

Hiver 2020 - COVID-19 quarantaine # 0

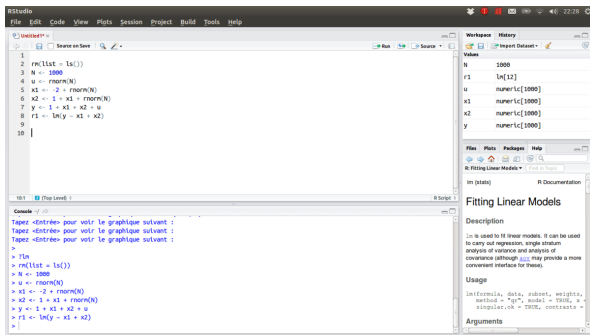


## Trois séries de capsules vidéos

- **régression Bernoulli**  $y_i \in \{0, 1\}$  ou  $Y_i \sim \mathcal{B}(p_i)$   
et extensions:  $y_i \in \{0, 1, \dots, n\}$  ou  $Y_i \sim \mathcal{B}(n, p_i)$  (binomiale)  
 $y_i \in \{A, B, C\}$  ou  $\mathbf{Y} = (\mathbf{1}_A, \mathbf{1}_B, \mathbf{1}_C) \sim \mathcal{M}(\mathbf{p})$  (multinomiale)  
#1 - #7
- **régression de Poisson**  $y_i \in \mathbb{N}$  ou  $Y_i \sim \mathcal{P}(\lambda_i)$ ,  $\lambda_i = \exp(\mathbf{x}_i^\top \boldsymbol{\beta})$   
#8 - #11
- **modèles GLM**  $Y_i \sim \mathcal{L}(\theta_i)$  avec  $\mathbb{E}(y_i) = g^{-1}(\mathbf{x}_i^\top \boldsymbol{\beta})$
- famille exponentielle, inférence, déviance, surdispersion, tests  
#12 - #19

# GLM (pratique)

Mise en oeuvre sous R (<https://rstudio.cloud/> version en ligne)



The screenshot shows the RStudio environment. The script editor contains the following R code:

```
1 rm(list = ls())
2 N <- 1000
3 u <- rnorm(N)
4 x1 <- -2 + rnorm(N)
5 x2 <- 1 + x1 + rnorm(N)
6 y <- 1 + x1 + x2 + u
7 r1 <- ln(y - x1 + x2)
8
9
10
```

The console shows the execution of the code with the following output:

```
> rm(list = ls())
> N <- 1000
> u <- rnorm(N)
> x1 <- -2 + rnorm(N)
> x2 <- 1 + x1 + rnorm(N)
> y <- 1 + x1 + x2 + u
> r1 <- ln(y - x1 + x2)
>
```

The workspace on the right shows the following objects:

Object	Class
N	numeric[1000]
u	numeric[1000]
x1	numeric[1000]
x2	numeric[1000]
y	numeric[1000]

The R Documentation pane on the right shows the documentation for the `lm` function, titled "Fitting Linear Models".

```
1 > ?glm
2 > reg = glm(y ~ x1+x2, family=binomial(link="logit"))
```

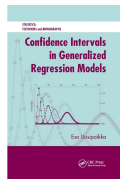
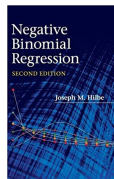
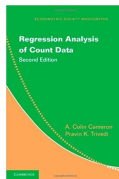
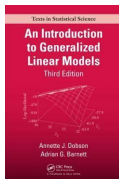
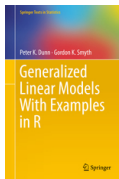
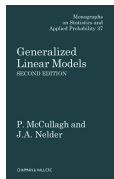
Complément : [github/STT5100](https://github.com/STT5100) ou [github/MAT7381](https://github.com/MAT7381), et [blog](#)  
Le rendu des devoirs se fera via [moodle](#)

# GLM (pratique)

Pour le cours [STT5100](#), les devoirs sont en ligne sur github,

```
1 > code_permanent = "ABCD12345678"  
2 > loc_fichier = paste("http://freakonometrics.free.fr/  
   /",code_permanent,"H2020D3.RData",sep="")  
3 > download.file(loc_fichier, "base_devoir_3.RData")  
4 > load("base_devoir_3.RData")  
5 > dim(database)  
6 [1] 6 24
```

(le code permanent est à changer...)



Pour toutes suggestions [arthur.charpentier@gmail.com](mailto:arthur.charpentier@gmail.com)

