

Modèles Linéaires Appliqués

Arthur Charpentier

Automne 2020

Rappels #0 (R, RStudio & markdown)

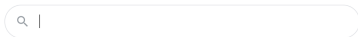
Dealing with errors, in R

```
1 > a=1
2 > A
3 Error: object 'A' not found
```

1) read the message







```
1 > code_permanent = "ABCD12345678"
2 > loc_fichier = paste("http://freakonometrics.free.fr/",code_permanent,"H2020D1.
  RData",sep="")
3 > download.file(loc_fichier, "base_devoir_1.RData")
4 > load("base_devoir_1.RData")
5 > step(lm(TARGET_deathRate~., data=database))
6 Error in step(lm(TARGET_deathRate ~ ., data = database)) :
7   AIC is -infinity for this model, so 'step' cannot proceed
```

2) use google...



Dealing with errors, in R

"AIC is -infinity for this model, so 'step' cannot proceed"

 Tous  Vidéos  Actualités  Shopping  Maps  Plus Paramètres

Environ 200 résultats (0,39 secondes)

stackoverflow.com › questions › ste... [Traduire cette page](#)

stepAIC error: AIC is -infinity - Stack Overflow

28 avr. 2018 - ... selection and receive an error: "AIC is -infinity for this model, so 'step' cannot proceed" when including all candidate predictor variables.

I'd expect that if you had a perfect fit - no errors. Make sure you don't have a copy of your dependent variable in with your predictors. If you need more help, you'll probably need to post a reproducible example.

– [Gregor Thomas](#) Apr 27 '18 at 18:57

stats.stackexchange.com › questions [Traduire cette page](#)

How to perform step() when $n < p$ in R? - Cross Validated

Trying to use step() for such problems i get the error message **AIC is -infinity for this model, so 'step' cannot proceed** . Is there a way to modify parameters so i ...

1 réponse

```
1 > dim(database)
2 [1] 6 32
```

3) ask me, or the teaching assistant, or post a question on the forum...

R + Markdown

R: ACT3035 - Laboratoire d'actuariat, *Logiciel R : types, sélection, entrée/sortie, graphisme, programmation de base*

- <http://beginr.u-bordeaux.fr/>
- <https://statistique-et-logiciel-r.com/>
- <http://www.jacolienvanrij.com/Tutorials/>
- <https://rmarkdown.rstudio.com/>
- <http://www.unexpected-vortices.com/sw/rippledoc/>
- <https://r4ds.had.co.nz/r-markdown.html>
- <https://bookdown.org/yihui/rmarkdown/>

Create a Rmd file

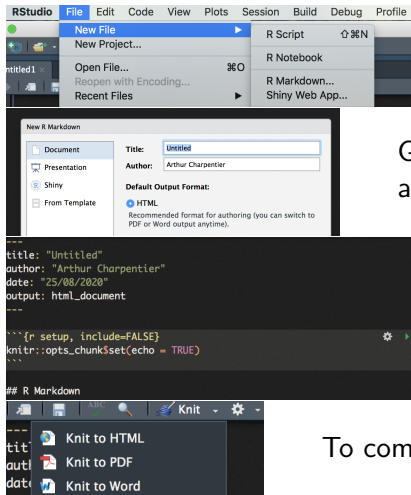
File

- > New File
- > R Markdown

Give a name for the document
ask for a **html** output

You get a template file,
with basic instructions

To compile, use **knit** (to html)



Preamble

Devoir 1 - STT5100

John Doe (ABCD12345678)

September 1st, 2020

Introduction

Nous allons commencer par importer les données

```
code_permanent = "ABCD12345678"
loc_fichier = paste("http://freakonometrics.free.fr/",code_permanent,"H202001.Rdata",sep="")
download.file(loc_fichier,"base_devoir_1.Rdata")
load("base_devoir_1.Rdata")
dim(database)
```

```
## [1] 4 32
```

Introduction

Premier modèle

Second modèle

Analyse

Devoir 1 - STT5100

John Doe (ABCD12345678)

September 1st, 2020

Introduction

Nous allons commencer par importer les données

```
code_permanent = "ABCD12345678"
loc_fichier = paste("http://freakonometrics.free.fr/",code_permanent,"H202001.Rdata",sep="")
download.file(loc_fichier,"base_devoir_1.Rdata")
load("base_devoir_1.Rdata")
dim(database)
```

```
## [1] 4 32
```

Devoir 1 - STT5100

John Doe (ABCD12345678)

September 1st, 2020

Introduction

Nous allons commencer par importer les données

```
code_permanent = "ABCD12345678"
loc_fichier = paste("http://freakonometrics.free.fr/",code_permanent,"H202001.Rdata",sep="")
download.file(loc_fichier,"base_devoir_1.Rdata")
load("base_devoir_1.Rdata")
dim(database)
```

```
## [1] 4 32
```

Devoir 1 - STT5100

John Doe (ABCD12345678)

September 1st, 2020

Introduction

Nous allons commencer par importer les données

```
code_permanent = "ABCD12345678"
loc_fichier = paste("http://freakonometrics.free.fr/",code_permanent,"H202001.Rdata",sep="")
download.file(loc_fichier,"base_devoir_1.Rdata")
load("base_devoir_1.Rdata")
dim(database)
```

```
## [1] 4 32
```

```
1 ---
2 title: "Devoir 1 - STT5100"
3 author: "John Doe (ABCD12345678)"
4 date: September 1st, 2020
5 output: html_document
6 ---
```

```
1 ---
2 title: "STT5100"
3 output:
4   html_document:
5     toc: true
6     toc_float: true
7 ---
```

```
1 ---
2 title: "STT5100"
3 output:
4   html_document:
5     theme: united
6     highlight: tango
7 ---
```

```
1 ---
2 title: "STT5100"
3 output:
4   html_document:
5     theme: journal
6     highlight: pygments
7 ---
```

Preamble

Devoir 1 - STT5100

John Doe (ABCD12345678)

September 1st, 2020

Introduction

Nous allons commencer par importer les données

```
code_permanent = "AKD12345678"
loc_fichier = paste("http://freakonometrics.free.fr/",code_permanent,"R02001.Rdata",sep="")
download.file(loc_fichier,"base_devoir_1.Rdata")
load("base_devoir_1.Rdata")
dim(database)
```

```
## [1] 6 32
```

Devoir 1 - STT5100

John Doe (ABCD12345678)

September 1st, 2020

Introduction

Nous allons commencer par importer les données

```
code_permanent = "AKD12345678"
loc_fichier = paste("http://freakonometrics.free.fr/",code_permanent,"R02001.Rdata",sep="")
download.file(loc_fichier,"base_devoir_1.Rdata")
load("base_devoir_1.Rdata")
dim(database)
```

```
## [1] 6 32
```

Devoir 1 - STT5100

John Doe (ABCD12345678)

September 1st, 2020

Introduction

Nous allons commencer par importer les données

```
code_permanent = "AKD12345678"
loc_fichier = paste("http://freakonometrics.free.fr/",code_permanent,"R02001.Rdata",sep="")
download.file(loc_fichier,"base_devoir_1.Rdata")
load("base_devoir_1.Rdata")
dim(database)
```

```
## [1] 6 32
```

Devoir 1 - STT5100

John Doe (ABCD12345678)

September 1st, 2020

Introduction

Nous allons commencer par importer les données

```
code_permanent = "AKD12345678"
loc_fichier = paste("http://freakonometrics.free.fr/",code_permanent,"R02001.Rdata",sep="")
download.file(loc_fichier,"base_devoir_1.Rdata")
load("base_devoir_1.Rdata")
dim(database)
```

```
## [1] 6 32
```

```
1 ---
2 title: "STT5100"
3 output:
4   html_document:
5     theme: sandstone
6     highlight: espresso
7 ---
```

```
1 ---
2 title: "STT5100"
3 output:
4   html_document:
5     theme: spacelab
6     highlight: kate
7 ---
```

```
1 ---
2 title: "STT5100"
3 output:
4   html_document:
5     theme: darkly
6     highlight: zenburn
7 ---
```

```
1 ---
2 title: "STT5100"
3 output:
4   html_document:
5     theme: cerulean
6 ---
```

Introduction

Sous-section

- la variable `y`
- la variable `x1`
- la variable `x2`

On peut faire des calculs et insérer les résultats dans le texte,

```
a = 2
b = 4
```

La variable `a` vaut 2 et la somme de `a` et `b` 6

et insérer des formules \LaTeX

Notre modèle est

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \beta_3 (x_1 - 10)_+^2 + \epsilon_i$$

et choisir la précision des nombres

On peut choisir la précision d'affichage des *nombres* dans le texte

```
a = 2
b = 4
(a+b)/7
```

```
## [1] 0.8571429
```

On peut afficher 0.8571429 ou 0.857

```
1 # Introduction
2
3 ## Sous-section
4
5 - la variable 'y'
6 - la variable 'x1'
7 - la variable 'x2'
```

```
1 '{r}
2 a = 2
3 b = 4
4 ''
5
6 La variable 'a' vaut 'r a' et la
   somme de 'a' et 'b' 'r a+b'
```

```
1 Notre modele est $$y_i=\beta_0 +
2 \beta_1x_1 + \beta_2x_1^2 +
3 \beta_3(x_1-10)_+^2 + \epsilon_i$$
```

```
1 On peut **choisir** la precision d'
   affichage des *nombres* dans le
   texte
2
3 On peut afficher 'r (a+b)/7' ou 'r
   format((a+b)/7,, digits=3)'
```


Syntax

```
model = step(lm(TARGET_deathRate~.,data=database[,1:5]))
```

```
## Start: AIC=21.32
## TARGET_deathRate ~ avgAnnCount + avgDeathsPerYear + incidenceRate
## medIncome
##
##           Df Sum of Sq  RSS   AIC
## <none>                 39.58 21.320
## - incidenceRate      1    91.76 131.35 26.516
## - avgAnnCount         1   471.20 510.79 34.665
## - avgDeathsPerYear    1   471.57 511.16 34.670
## - medIncome           1   491.41 531.00 34.898
```

```
AIC(model)
```

```
## [1] 40.3473
```

```
model = step(lm(TARGET_deathRate~.,data=database[,1:5]),trace=0)
AIC(model)
```

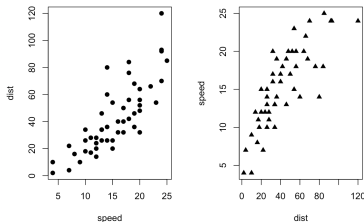
```
## [1] 40.3473
```

```
1  "{r}
2  model = step(lm(TARGET_deathRate~.,
3  AIC(model)
4  "
5
6  "{r}
7  model = step(lm(TARGET_deathRate~.,
8  AIC(model)
9  "

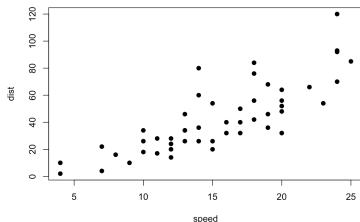
```

Graphics

```
par(mfrow=c(1,2))
plot(cars, pch = 19)
plot(cars[,2:1], pch = 17)
```



```
plot(cars, pch = 19)
```

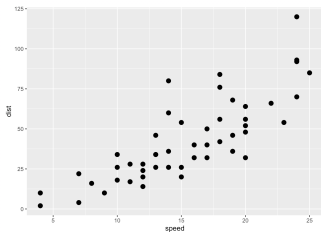


```
1 '{{{r}
2 par(mfrow=c(1,2))
3 plot(cars, pch = 19)
4 plot(cars[,2:1], pch = 17)
5 ''''
```

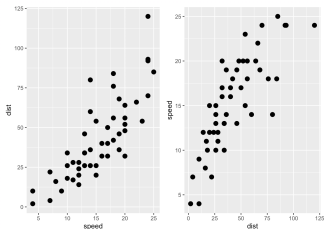
```
1 '{{{r}
2 plot(cars, pch = 19)
3 ''''
```

Graphics

```
library(ggplot2)
ggplot(cars, aes(x=speed, y=dist)) +
  geom_point(size=3)
```



```
library(gridExtra)
p1 = ggplot(cars, aes(x=speed, y=dist)) +
  geom_point(size=3)
p2 = ggplot(cars, aes(x=dist, y=speed)) +
  geom_point(size=3)
grid.arrange(p1, p2, ncol = 2)
```



or, with ggplot2 graphs

```
1 {{r}
2 library(ggplot2)
3 ggplot(cars, aes(x=speed, y=dist)) +
4   geom_point(size=3)
5 ""
```

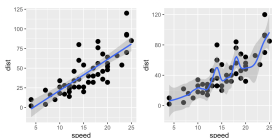
```
1 {{r}
2 library(gridExtra)
3 p1 = ggplot(cars, aes(x=speed, y=dist
4   )) +
5   geom_point(size=3)
6 p2 = ggplot(cars, aes(x=dist, y=speed
7   )) +
8   geom_point(size=3)
9 grid.arrange(p1, p2, ncol = 2)
10 ""
```

Graphics

```
library(gridExtra)
p1 = ggplot(cars, aes(x=speed, y=dist)) +
  geom_point(size=3) +
  geom_smooth(method='lm')
p2 = ggplot(cars, aes(x=speed, y=dist)) +
  geom_point(size=3) +
  geom_smooth(span = 0.3)
grid.arrange(p1,p2, ncol = 2)
```

```
## 'geom_smooth()' using formula 'y ~ x'
```

```
## 'geom_smooth()' using method = 'loess' and formula 'y ~ x'
```



we can also set width/height

```
1  “{r fig.width=6, fig.height=3}
2  library(gridExtra)
3  p1 = ggplot(cars, aes(x=speed, y=dist
4    )) +
5    geom_point(size=3)+
6    geom_smooth(method='lm')
7  p2 = ggplot(cars, aes(x=speed, y=dist
8    )) +
9    geom_point(size=3)+
10   geom_smooth(span = 0.3)
11   grid.arrange(p1,p2, ncol = 2)
12   “
```

Tables

```
tableau = data.frame(model = 1:4, AIC=rexp(4), R2=runif(4))
knitr::kable(tableau, caption = 'Comparison')
```

Comparison

model	AIC	R2
1	0.1116390	0.7006295
2	0.5923306	0.1299259
3	0.8882854	0.7452766
4	0.7750341	0.2459111

```
library(kableExtra)
knitr::kable(tableau, caption = 'Comparison') %>%
  kable_styling(bootstrap_options = c("striped", "hover"))
```

Comparison

model	AIC	R2
1	0.1116390	0.7006295
2	0.5923306	0.1299259
3	0.8882854	0.7452766
4	0.7750341	0.2459111

```
1 "{r}"
2 tableau = data.frame(model = 1:4, AIC
3   =rexp(4), R2=runif(4))
4 knitr::kable(tableau, caption = '
5   Comparison')
```

```
1 "{r}"
2 library(kableExtra)
3 knitr::kable(tableau, caption = '
4   Comparison') %>%
5 kable_styling(bootstrap_options = c("
6   striped", "hover"))
```