

Régression linéaire simple

Timothée Flutre

26/01/2015

Contents

1	Préambule	1
2	Introduction	2
3	Définir la question	3
4	Ecrire le modèle	3
4.1	Notations	3
4.2	Vraisemblance	3
5	Simuler des données	3
6	Réaliser l'inférence	3
7	Explorer les simulations possibles	3
8	Appendix	4

1 Préambule

Ce document a été généré à partir d'un fichier texte au format Rmd utilisé avec le logiciel libre [R](#). Pour exporter un tel fichier vers les formats HTML et PDF, installez le paquet [rmarkdown](#) disponible sur CRAN (il va vraisemblablement vous être demandé d'installer d'autres paquets), puis ouvrez R et entrez:

```
library(rmarkdown)
render("myanalysis.Rmd", "all")
```

Il est généralement plus simple d'utiliser le logiciel libre [RStudio](#), mais ce n'est pas obligatoire.

Pour plus de détails, lisez [cette page](#), et regardez les exemples ci-dessous.

Exemple de liste non ordonnée:

- aaa
- bbb

Exemple de liste ordonnée:

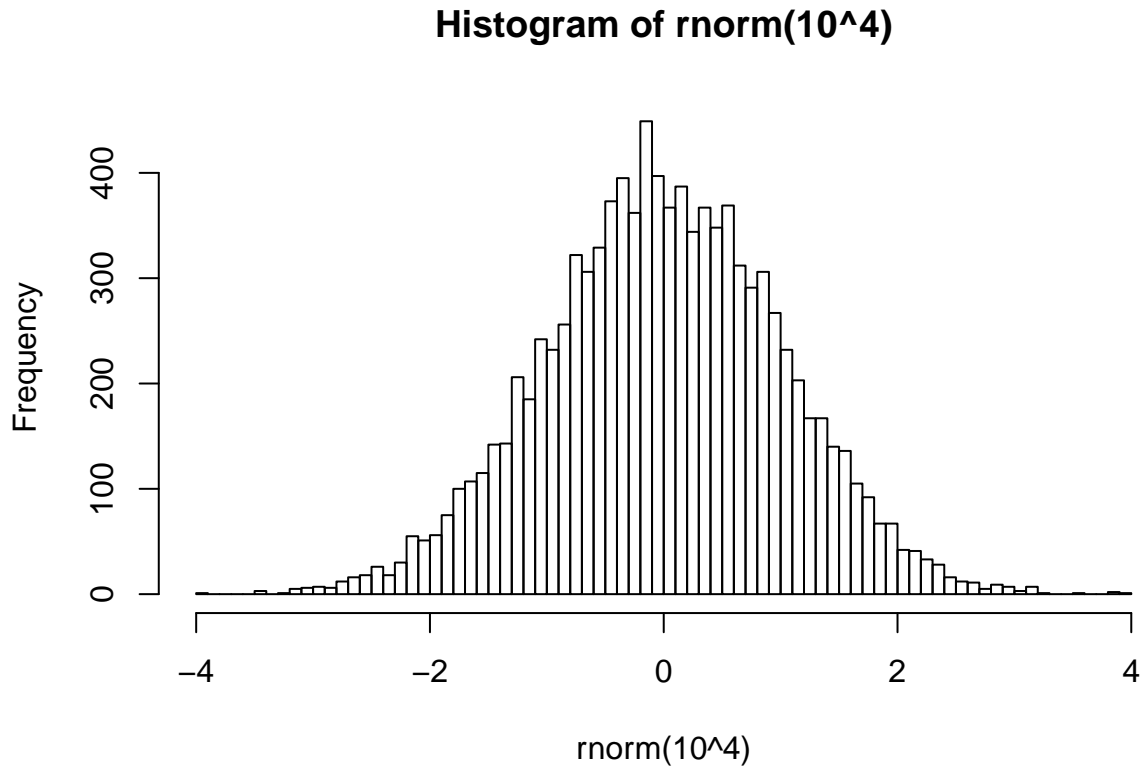
1. ccc

2. ddd

Exemple de mot en **gras** ou *italique*.

Exemple de graphique à partir de R:

```
hist(rnorm(10^4), breaks="FD")
```



Exemple de tableau à partir de R:

```
knitr::kable(mtcars[1:5,])
```

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
Mazda RX4	21.0	6	160	110	3.90	2.62	16.5	0	1	4	4
Mazda RX4 Wag	21.0	6	160	110	3.90	2.88	17.0	0	1	4	4
Datsun 710	22.8	4	108	93	3.85	2.32	18.6	1	1	4	1
Hornet 4 Drive	21.4	6	258	110	3.08	3.21	19.4	1	0	3	1
Hornet Sportabout	18.7	8	360	175	3.15	3.44	17.0	0	0	3	2

Exemple d'équation écrite en [LaTeX](#) ([livre en ligne](#)): $y_i \sim \mathcal{N}(\mu, \sigma^2)$

2 Introduction

Ce document fait partie de l'atelier "Prédiction Génomique" organisé et animé par Jacques David et Timothée Flutre en 2015 à Montpellier SupAgro dans le cadre du master APIMET.

Ce document a pour but d'introduire les étudiants à la simulation en prenant comme exemple la régression linéaire simple.

3 Définir la question

...

4 Ecrire le modèle

4.1 Notations

- n : nombre d'individus
- i : indice indiquant le i -ème individu, donc $i \in \{1, \dots, n\}$
- y_i : phénotype de l'individu i
- μ : moyenne globale du phénotypes des n individus
- x_i : génotype de l'individu i au marqueur SNP d'intérêt ($\in \{0, 1, 2\}$)
- f : fréquence de l'allèle minoritaire à ce SNP
- β : effet additif de chaque copie de l'allèle minoritaire en unité du phénotype
- ϵ_i : erreur pour l'individu i
- σ^2 : variance des erreurs

4.2 Vraisemblance

$$\forall i \quad y_i = \mu + \beta x_i + \epsilon_i \text{ avec } \epsilon_i \sim \mathcal{N}(0, \sigma^2)$$

5 Simuler des données

...

6 Réaliser l'inférence

...

7 Explorer les simulations possibles

...

8 Appendix

```
print(sessionInfo(), locale=FALSE)
```

```
## R version 3.1.2 (2014-10-31)
## Platform: x86_64-pc-linux-gnu (64-bit)
##
## attached base packages:
## [1] stats      graphics  grDevices  utils      datasets  methods   base
##
## other attached packages:
## [1] MASS_7.3-37      knitr_1.9        rmarkdown_0.4.2
##
## loaded via a namespace (and not attached):
## [1] digest_0.6.8     evaluate_0.5.5   formatR_1.0      htmltools_0.2.6
## [5] stringr_0.6.2    tools_3.1.2      yaml_2.1.13
```