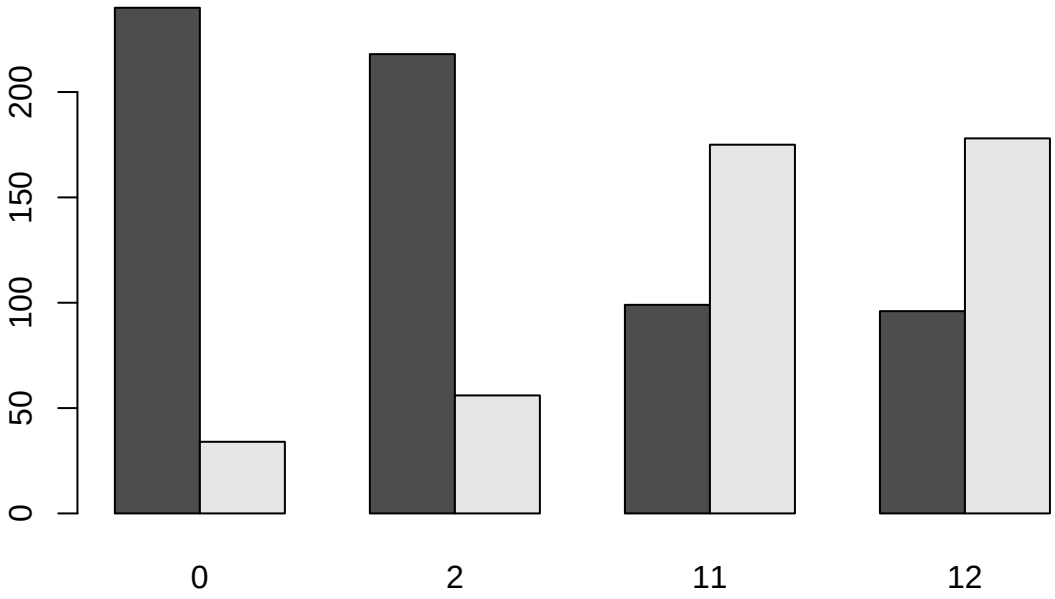
Calcul SE simulations Marilou

Setting up the system

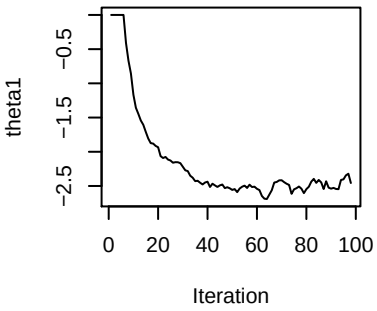
Binomial model in saemix

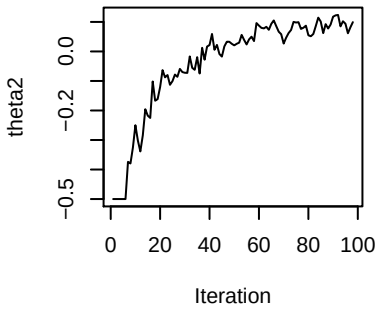
```
##
##
## The following SaemixModel object was successfully created:
##
## Nonlinear mixed-effects model
##   Model function:  Binary model  Model type:  likelihood
## function(psi,id,xidep) {
##   tim<-xidep[,1]
##   y<-xidep[,2]
##   inter<-psi[id,1]
##   slope<-psi[id,2]
##   logit<-inter+slope*tim
##   pevent<-exp(logit)/(1+exp(logit))
##   logpdf<-rep(0,length(tim))
##   P.obs = (y==0)*(1-pevent)+(y==1)*pevent
##   logpdf <- log(P.obs)
##   return(logpdf)
## }
##   Nb of parameters: 2
##       parameter names:  theta1 theta2
##       distribution:
##      Parameter Distribution Estimated
## [1,] theta1    normal       Estimated
## [2,] theta2    normal       Estimated
##   Variance-covariance matrix:
##        theta1 theta2
## theta1      1      0
## theta2      0      1
##   Covariate model:
##      theta1 theta2
## [1,]      0      1
##     Initial values
##              theta1 theta2
## Pop.CondInit      0   -0.5
## Cov.CondInit      0    0.5
```

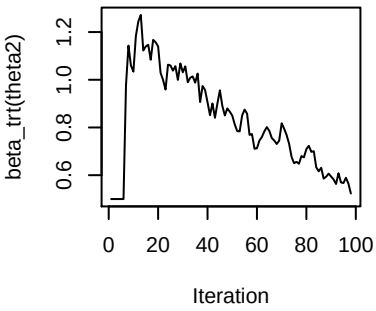
Run saemix sur le premier jeu de données pour vérifier que ça colle

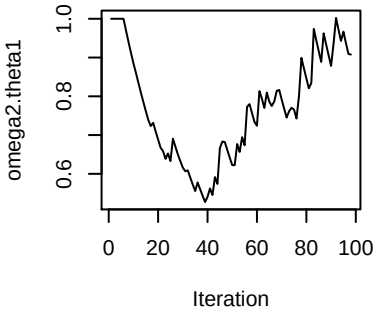


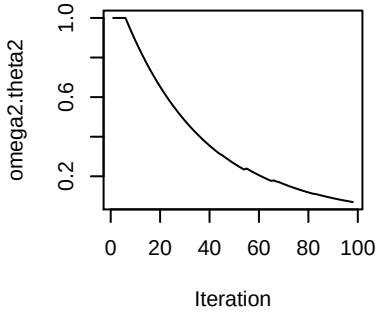

```
## Reading data from file /home/eco/work/theses/sirM2marilou/simulations/data/data1.txt
## These are the first lines of the dataset as read into R. Please check the format of the data is appro
##     id    theta1     theta2 betacov trt time    logitp          p y
## 1    1 -3.096707 0.16248408    0.45   0    0 -3.096707 0.04324328 0
## 275  1 -3.096707 0.16248408    0.45   0    2 -2.771739 0.05887058 0
## 549  1 -3.096707 0.16248408    0.45   0   11 -1.309382 0.21259020 0
## 823  1 -3.096707 0.16248408    0.45   0   12 -1.146898 0.24105606 0
## 2    2 -2.821699 0.06527274    0.45   0    0 -2.821699 0.05616282 0
## 276  2 -2.821699 0.06527274    0.45   0    2 -2.691153 0.06349740 0
## Column name(s)  do(es) not exist in the dataset, please check
## Remove columns 1 (  )
## No valid name given, attempting automatic recognition
## Automatic recognition of columns y successful
## [1] "trt"
##
##
## The following SaemixData object was successfully created:
##
## Object of class SaemixData
##     longitudinal data for use with the SAEM algorithm
## Dataset /home/eco/work/theses/sirM2marilou/simulations/data/data1.txt
##     Structured data: y ~ time + y | id
##     X variable for graphs: time ()
##     covariates: trt (-)
##       reference class for covariate trt :  0

## Running main SAEM algorithm
## [1] "Thu Jul  4 22:56:56 2019"
## .
```


theta1
−2.5
−1.5
−0.5
0
20
40
60
80
100
Iteration

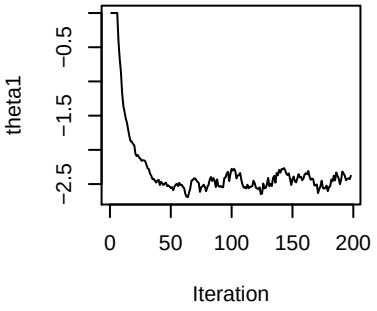

theta2
−0.5
−0.2
0.0
0
20
40
60
80
100
Iteration

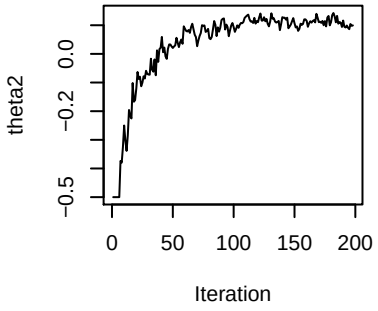

beta_trt(theta2)
0.6
0.8
1.0
1.2
0
20
40
60
80
100
Iteration

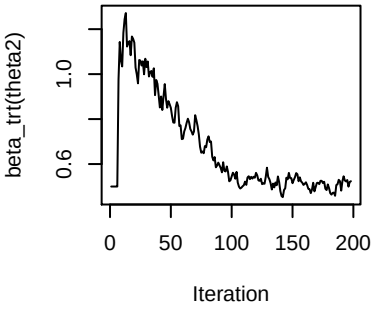

omega2.theta1
0.6
0.8
1.0
0
20
40
60
80
100
Iteration

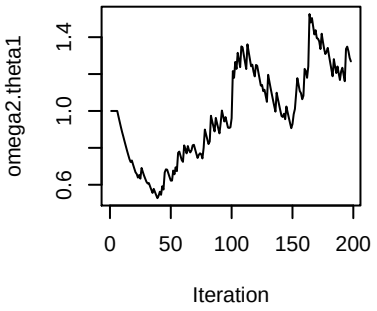

omega2.theta2
0.2
0.6
1.0
0
20
40
60
80
100
Iteration

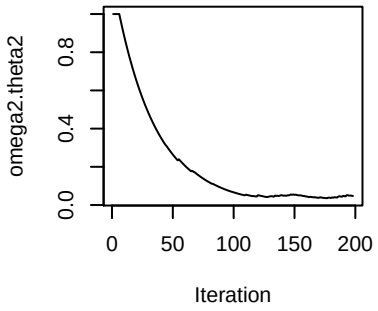
.


theta1
−2.5
−1.5
−0.5
0
50
100
150
200
Iteration

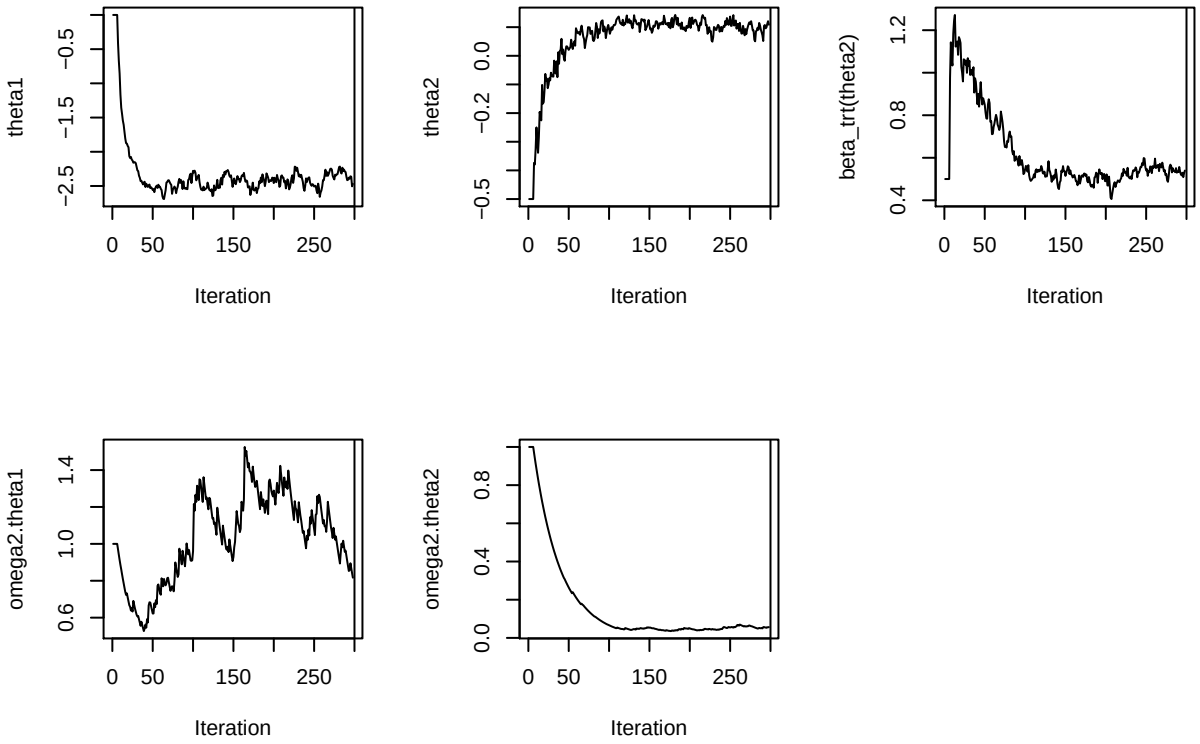

theta2
−0.5
−0.2
0.0
0
50
100
150
200
Iteration


beta_trt(theta2)
0.6
1.0
0
50
100
150
200
Iteration

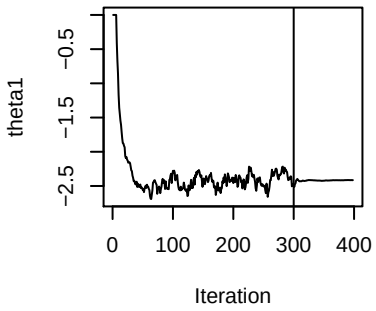

omega2.theta1
0.6
1.0
1.4
0
50
100
150
200
Iteration

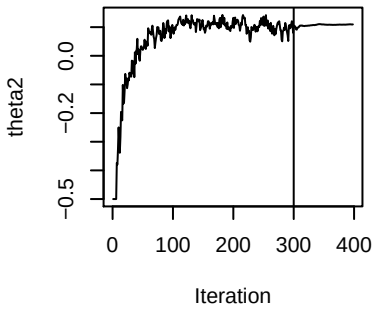

omega2.theta2
0.0
0.4
0.8
0
50
100
150
200
Iteration

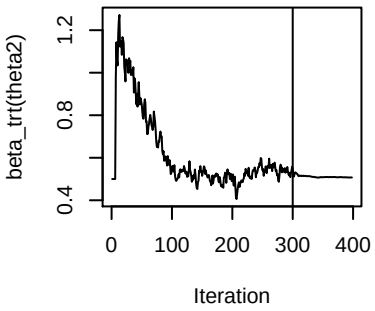
.

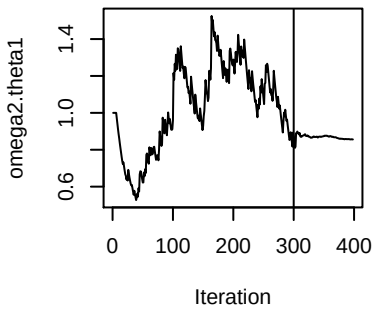


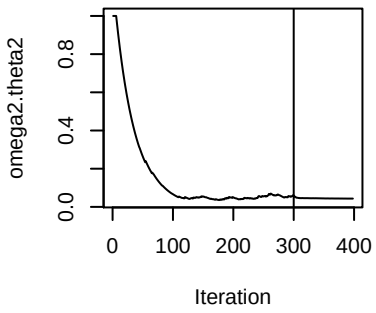

```
## .
##
##     Minimisation finished
## [1] "Thu Jul  4 22:56:59 2019"
## Error in solve.default(FO) :
##   le système est numériquement singulier : conditionnement de la réciproque = 2.99566e-17
## Error computing the Fisher Information Matrix: singular system.
```

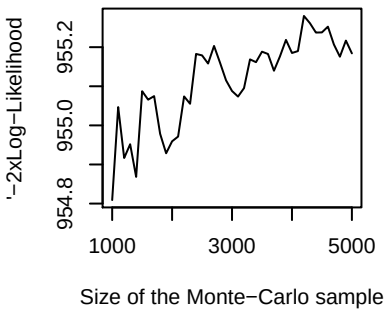













```
## Nonlinear mixed-effects model fit by the SAEM algorithm
## ----------------------------------
## ----          Data              ----
## ----------------------------------
## Object of class SaemixData
##     longitudinal data for use with the SAEM algorithm
## Dataset /home/eco/work/theses/sirM2marilou/simulations/data/data1.txt
##     Structured data: y ~ time + y | id
##     X variable for graphs: time ()
##     covariates: trt (-)
##       reference class for covariate trt :  0
## Dataset characteristics:
##     number of subjects:     274
##     number of observations: 1096
##     average/min/max nb obs: 4.00  /  4  /  4
## First 10 lines of data:
##     id time y y.1 trt mdv cens occ ytype
## 1    1    0 0   0   0   0    0   1     1
## 275  1    2 0   0   0   0    0   1     1
## 549  1   11 0   0   0   0    0   1     1
## 823  1   12 0   0   0   0    0   1     1
## 2    2    0 0   0   0   0    0   1     1
## 276  2    2 0   0   0   0    0   1     1
## 550  2   11 0   0   0   0    0   1     1
## 824  2   12 0   0   0   0    0   1     1
## 3    3    0 0   0   0   0    0   1     1
## 277  3    2 0   0   0   0    0   1     1
## ----------------------------------
## ----          Model             ----
## ----------------------------------
```

```
## Nonlinear mixed-effects model
##   Model function:  Binary model  Model type:  likelihood
## function(psi,id,xidep) {
##   tim<-xidep[,1]
##   y<-xidep[,2]
##   inter<-psi[id,1]
##   slope<-psi[id,2]
##   logit<-inter+slope*tim
##   pevent<-exp(logit)/(1+exp(logit))
##   logpdf<-rep(0,length(tim))
##   P.obs = (y==0)*(1-pevent)+(y==1)*pevent
##   logpdf <- log(P.obs)
##   return(logpdf)
## }
## <bytecode: 0x6bf5b18>
##   Nb of parameters: 2
##       parameter names:  theta1 theta2
##       distribution:
##      Parameter Distribution Estimated
## [1,] theta1    normal       Estimated
## [2,] theta2    normal       Estimated
##   Variance-covariance matrix:
##        theta1 theta2
## theta1      1      0
## theta2      0      1
##   Covariate model:
##     [,1] [,2]
## trt    0    1
##     Initial values
##              theta1 theta2
## Pop.CondInit      0   -0.5
## Cov.CondInit      0    0.5
## ----------------------------------
## ----    Key algorithm options  ----
## ----------------------------------
##     Estimation of individual parameters (MAP)
##     Estimation of standard errors and linearised log-likelihood
##     Estimation of log-likelihood by importance sampling
##     Number of iterations:  K1=300, K2=100
##     Number of chains:  1
##     Seed:  1234567
##     Number of MCMC iterations for IS:  5000
##     Simulations:
##         nb of simulated datasets used for npde:  1000
##         nb of simulated datasets used for VPC:  100
##     Input/output
##         save the results to a file:  FALSE
##         save the graphs to files:  FALSE
## ----------------------------------------------------
## ----                  Results                   ----
## ----------------------------------------------------
## -----------------  Fixed effects  ------------------
## ----------------------------------------------------
##      Parameter        Estimate SE   CV(%) p-value
```

```
## [1,] theta1           -2.42    0.21  8.5  -
## [2,] theta2            0.11    0.10 95.2  -
## [3,] beta_trt(theta2)  0.51    0.18 36.1  0.0028
## ----------------------------------------------------
## -----------  Variance of random effects  -----------
## ----------------------------------------------------
##        Parameter     Estimate SE CV(%)
## theta1 omega2.theta1 0.856    NA NA
## theta2 omega2.theta2 0.043    NA NA
## ----------------------------------------------------
## ------  Correlation matrix of random effects  ------
## ----------------------------------------------------
##               omega2.theta1 omega2.theta2
## omega2.theta1 1             0
## omega2.theta2 0             1
## ----------------------------------------------------
## ---------------  Statistical criteria  -------------
## ----------------------------------------------------
## Likelihood computed by linearisation
##       -2LL= 3252.623
##       AIC = 3264.623
##       BIC = 3286.301
##
## Likelihood computed by importance sampling
##       -2LL= 955.1842
##       AIC = 967.1842
##       BIC = 988.8629
## ----------------------------------------------------
```

Run saemix sur les 200 simulations de Marilou, et récupération des résultats dans un tableau, fait dans computeSEemp_binomial.R

Calcul SE par bootstrap pour 1 jeu de données

Test rapide avec 10 échantillons bootstrap seulement, juste pour tester les codes

```
## Warning in rm(.Random.seed): objet '.Random.seed' introuvable

## Warning in rm(.Random.seed): objet '.Random.seed' introuvable

## Warning in rm(.Random.seed): objet '.Random.seed' introuvable

## Warning in rm(.Random.seed): objet '.Random.seed' introuvable

## Warning in rm(.Random.seed): objet '.Random.seed' introuvable

## Warning in rm(.Random.seed): objet '.Random.seed' introuvable

## Warning in rm(.Random.seed): objet '.Random.seed' introuvable

## Warning in rm(.Random.seed): objet '.Random.seed' introuvable

## Warning in rm(.Random.seed): objet '.Random.seed' introuvable
```

```
## Warning in rm(.Random.seed): objet '.Random.seed' introuvable

## Warning in rm(.Random.seed): objet '.Random.seed' introuvable

## Warning in rm(.Random.seed): objet '.Random.seed' introuvable

## Warning in rm(.Random.seed): objet '.Random.seed' introuvable

## Warning in rm(.Random.seed): objet '.Random.seed' introuvable

## Warning in rm(.Random.seed): objet '.Random.seed' introuvable

## Warning in rm(.Random.seed): objet '.Random.seed' introuvable

## Warning in rm(.Random.seed): objet '.Random.seed' introuvable

## Warning in rm(.Random.seed): objet '.Random.seed' introuvable

## Warning in rm(.Random.seed): objet '.Random.seed' introuvable

## Warning in rm(.Random.seed): objet '.Random.seed' introuvable
```

Summary, SD and quantiles of bootstrap distributions for the 2 fixed, effect, the covariate effect and the 2 variabilities.

```r
apply(res.boot[,2:6],2,summary)
```

```
##            theta1    theta2 beta_trt(theta2) omega2.theta1 omega2.theta2
## Min.    -2.879033 0.0809111        0.3693210     0.3391882    0.01386664
## 1st Qu. -2.688144 0.1086251        0.4437737     0.9001025    0.02448749
## Median  -2.544041 0.1212871        0.5163572     1.2770004    0.03805155
## Mean    -2.519576 0.1258515        0.5053262     1.2546404    0.03943791
## 3rd Qu. -2.350707 0.1466459        0.5660670     1.5937432    0.04766536
## Max.    -2.128179 0.1672035        0.6018643     2.1754994    0.07907715
```

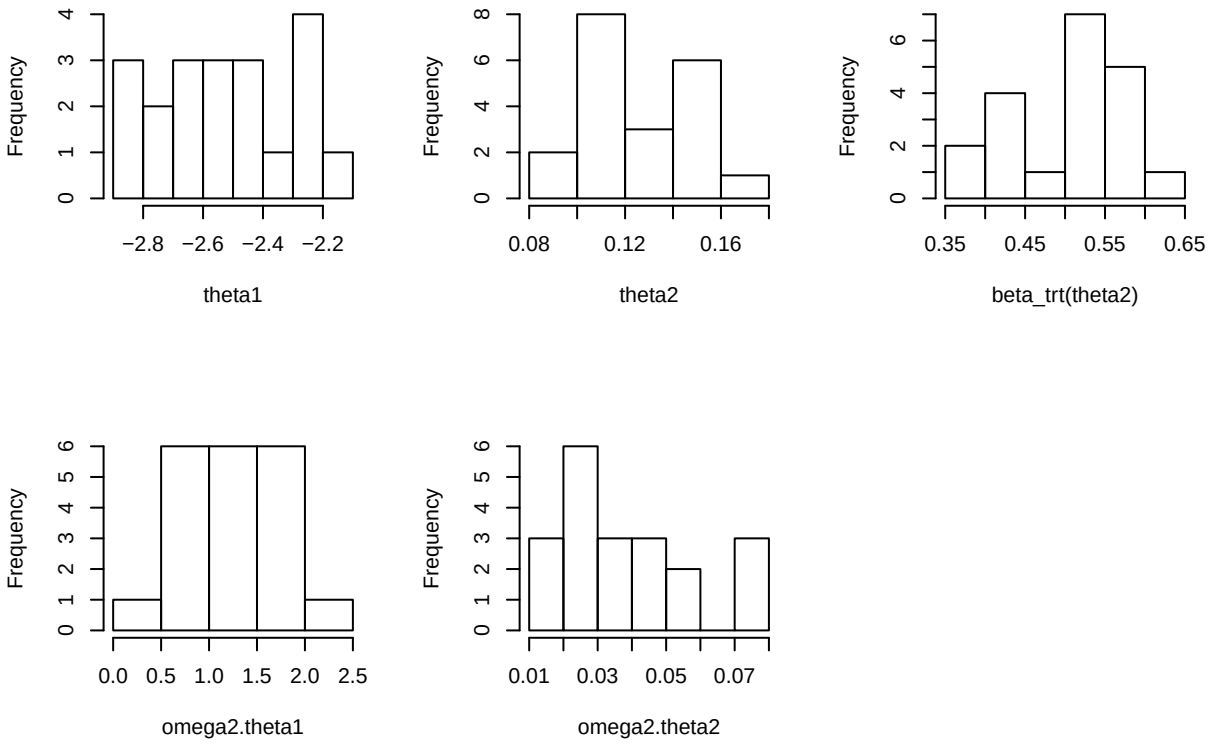
```r
apply(res.boot[,2:6],2,sd)
```

```
##           theta1           theta2 beta_trt(theta2)    omega2.theta1 
##       0.22495423       0.02378150       0.07213049       0.48297218 
##    omega2.theta2 
##       0.01976853
```

```r
apply(res.boot[,2:6],2,quantile,c(0.025,0.5,0.975))
```

```
##          theta1     theta2 beta_trt(theta2) omega2.theta1 omega2.theta2
## 2.5%  -2.873633 0.08529847        0.3784378     0.4890779    0.01587834
## 50%   -2.544041 0.12128714        0.5163572     1.2770004    0.03805155
## 97.5% -2.180187 0.16323033        0.5964226     2.0346047    0.07777023
```

Bootstrap distributions



Calcul SE par la AGQ (méthode de Sebastian, code ad hoc)

```
##                mu1        mu2         mu3   omega1.1    omega2.2
## mu1       59.72529  186.72123    77.07814  12.136375   -28.76754
## mu2      186.72123 2133.38367   761.53398  -7.881850  -881.84773
## mu3       77.07814  761.53398   761.53398 -10.076503 -1797.41995
## omega1.1  12.13637   -7.88185   -10.07650   7.735019    77.02905
## omega2.2 -28.76754 -881.84773 -1797.41995  77.029050  8663.43890
```

```
##       mu1       mu2       mu3  omega1.1  omega2.2
## -9.039789 30.771085 13.426490 62.569823 39.293465
```

```
##               param         se
## mu1      -2.4160841 0.21840891
## mu2       0.1102140 0.03391403
## mu3       0.5063923 0.06799071
## omega1.1  0.8562196 0.53573506
## omega2.2  0.0433050 0.01701604
```

Calcul SE empirique simulations Marilou

```
##         Simulation    theta1       theta2      beta omega2.theta1
## Min.          1.00 -2.788802 -0.009588763 0.3551997     0.3467233
## 1st Qu.      50.75 -2.330911  0.070810643 0.4763961     0.8026750
## Median      100.50 -2.200969  0.088915945 0.5244584     1.0636280
## Mean        100.50 -2.215034  0.090789687 0.5341374     1.1031808
## 3rd Qu.     150.25 -2.096785  0.110608475 0.5889836     1.3084848
## Max.        200.00 -1.729065  0.170173200 0.8098407     2.2509190
##         omega2.theta2
```

```
## Min.       0.01356832
## 1st Qu.    0.03139451
## Median     0.04040995
## Mean       0.04473278
## 3rd Qu.    0.05563287
## Max.       0.11319040
```

```
##    Simulation        theta1        theta2          beta omega2.theta1
##   57.87918451    0.19709055    0.02866786    0.07997905    0.41631827
## omega2.theta2
##    0.01855807
```

Comparaison des différentes estimations des SE

SE empiriques (sur 200 simulations), SE bootstrap (avec seulement 20 échantillons) et SE prédites par AGQ

```
## SE empiriques:
```

```
##    Simulation        theta1        theta2          beta omega2.theta1
##   57.87918451    0.19709055    0.02866786    0.07997905    0.41631827
## omega2.theta2
##    0.01855807
```

```
## SE bootstrap:
```

```
##           theta1           theta2 beta_trt(theta2)    omega2.theta1
##       0.22495423       0.02378150       0.07213049       0.48297218
##    omega2.theta2
##       0.01976853
```

```
## SE :
```

```
##        mu1        mu2        mu3   omega1.1   omega2.2
## 0.21840891 0.03391403 0.06799071 0.53573506 0.01701604
```

Calcul SE par bootstrap pour l'ensemble des jeux de données binomial

Test du bootstrap pour un modèle binomial (Eco TODO)

- Code bootstrap - en mode batch
- R CMD BATCH run_bootstrapBinomial.R