CHOIX DE MODELE, BAYES FACTOR

Quelques rappels ...
On veut choisir entre deux modèles M_1 et M_2

$$BF_{21} = [D|M_2]/[D|M_1] = \frac{\int [D|\theta, M_2][\theta|M_2]d\theta}{\int [D|\theta, M_1][\theta|M_1]d\theta}$$

où D sont les données.

On veut choisir entre K modèles M_1, ..., M_K
$\Rightarrow$ normalise les probabilités a posteriori des modèles

$$P(M_j|D) = \frac{BF_{ju}}{\sum_{k=1}^{K} BF_{ku}} = \frac{[D|M_j]}{\sum_{k=1}^{K} [D|M_k]}$$

quand équiprobabilité a priori entre les modèles.

CHOIX DE MODELE, BAYES FACTOR

Table "pratique" de Jeffreys pour la décision
en faveur de M_2 par rapport à M_1

- en termes de BF_{21}
- ou de probas a posteriori $P_{21} = \frac{[D|M_2]}{[D|M_1]+[D|M_2]}$

$$1 < BF_{21} < 3 \quad \Leftrightarrow \quad 0.5 < P_{21} < 0.75 \quad \Rightarrow \textit{faible}$$

$$3 < BF_{21} < 12 \quad \Leftrightarrow \quad 0.75 < P_{21} < 0.92 \quad \Rightarrow \textit{positif}$$

$$12 < BF_{21} < 150 \quad \Leftrightarrow \quad 0.92 < P_{21} < 0.99 \quad \Rightarrow \textit{fort}$$

$$BF_{21} > 150 \quad \Leftrightarrow \quad P_{21} > 0.99 \quad \Rightarrow \textit{decisif}$$

Approximations

Comment obtenir les BF_{21} si pas de calculs explicites ?

Comme $BF_{21} = [D|M_2]/[D|M_1]$,

$\Rightarrow$ Il faut approcher [D|M_j], $j = 1, 2$!

Idée 1 naive

Faire du Monte Carlo direct !

En effet,

$$[D|M_j] = \int [D|\theta, M_j][\theta|M_j]d\theta = E_{[\theta|M_j]}([D|\theta, M_j])$$

- Simuler $(\theta^1, \cdots \theta^T)$ sous la loi a priori $[\theta|M_j]$
- Approcher l'espérance par la vraisemblance moyenne

$$\widehat{[D|M_j]} = \frac{\sum [D|\theta^t, M_j]}{T}$$

Pas une bonne idée ...

Gros problèmes calculatoires surtout si priors vagues ! ! !

Idée 1 naive

Faire du Monte Carlo direct !

En effet,

$$[D|M_j] = \int [D|\theta, M_j][\theta|M_j]d\theta = E_{[\theta|M_j]}([D|\theta, M_j])$$

- Simuler $(\theta^1, \cdots \theta^T)$ sous la loi a priori $[\theta|M_j]$
- Approcher l'espérance par la vraisemblance moyenne

$$\widehat{[D|M_j]} = \frac{\sum [D|\theta^t, M_j]}{T}$$

Pas une bonne idée ...

Gros problèmes calculatoires surtout si priors vagues ! ! !

Idée 1 naive

Faire du Monte Carlo direct !

En effet,

$$[D|M_j] = \int [D|\theta, M_j][\theta|M_j] d\theta = E_{[\theta|M_j]}([D|\theta, M_j])$$

- Simuler $(\theta^1, \cdots \theta^T)$ sous la loi a priori $[\theta|M_j]$
- Approcher l'espérance par la vraisemblance moyenne

$$\widehat{[D|M_j]} = \frac{\sum [D|\theta^t, M_j]}{T}$$

Pas une bonne idée ...
Gros problèmes calculatoires surtout si priors vagues ! ! !

Idée 2 : Importance sampling

on montre facilement que (Bayes ...)

$$1/[D|M_j] = [\theta|D, M_j]/([D|\theta, M_j].[\theta|M_j])$$

donc, si g est une densité alors

$$1/[D|M_j] = 1/[D|M_j]\int g(\theta)d\theta = \int g(\theta)/([D|\theta, M_j].[\theta|M_j])[\theta|D, M_j]d\theta$$

- Simuler $(\theta^1, \cdots \theta^T)$ sous la loi $[\theta|D, M_j]$ (déjà avec MCMC !)
- Estimer $1/[D|M_j]$ par

$$1/\widehat{[D|M_j]} = \frac{\sum g(\theta^t)/([D|\theta^t, M_j].[\theta^t|M_j])}{T}$$

Idée 2 : Importance sampling

on montre facilement que (Bayes ...)

$$1/[D|M_j] = [\theta|D, M_j]/([D|\theta, M_j].[\theta|M_j])$$

donc, si g est une densité alors

$$1/[D|M_j] = 1/[D|M_j] \int g(\theta)d\theta = \int g(\theta)/([D|\theta, M_j].[\theta|M_j])[\theta|D, M_j]d\theta$$

- Simuler $(\theta^1, \cdots \theta^T)$ sous la loi $[\theta|D, M_j]$ (déjà avec MCMC !)
- Estimer $1/[D|M_j]$ par

$$1/\widehat{[D|M_j]} = \frac{\sum g(\theta^t)/([D|\theta^t, M_j].[\theta^t|M_j])}{T}$$

Idée 2 : Importance sampling

on montre facilement que (Bayes ...)

$$1/[D|M_j] = [\theta|D, M_j]/([D|\theta, M_j].[\theta|M_j])$$

donc, si g est une densité alors

$$1/[D|M_j] = 1/[D|M_j] \int g(\theta)d\theta = \int g(\theta)/([D|\theta, M_j].[\theta|M_j])[\theta|D, M_j]d\theta$$

- Simuler $(\theta^1, \cdots \theta^T)$ sous la loi $[\theta|D, M_j]$ (déjà avec MCMC !)
- Estimer $1/[D|M_j]$ par

$$1/\widehat{[D|M_j]} = \frac{\sum g(\theta^t)/([D|\theta^t, M_j].[\theta^t|M_j])}{T}$$

Idée 2 : Importance sampling

Choix de la densité g ?

Le plus usuel :

- g = priori mais pas une bonne idée à nouveau !
- g = densité d'une loi gaussienne (ou autre) de moyenne et variance raisonnables !

  $\Rightarrow$ Par ex, moy. et var a posteriori de θ estimées via MCMC

  ...

Idée 3 : Approximation de densités

On a : $[D|M_j] = ([D|\theta, M_j].[\theta|M_j])/[\theta|D, M_j]$
et ceci est vrai en particulier pour $\theta^* = E_{[\theta|D,M_j]}(\theta)$!

- Simuler $(\theta^1, \cdots \theta^T)$ sous la loi $[\theta|D, M_j]$ (MCMC !)
- Prendre un estimateur de $\theta^* \Rightarrow \theta^{**}{=}\sum \theta^t/T$
- Calcul de $([D|\theta^{**}, M_j].[\theta^{**}|M_j])$
- Estimation de la valeur de la densité h de $[\theta|D, M_j]$ en θ^{**} $\Rightarrow$ par ex, $\hat{h}(\theta^{**})$ estimateur a noyau gaussien (fonction density sous R) sur la base du MCMC

$$[D\hat{|}M_j] = ([D|\theta^{**}, M_j].[\theta^{**}|M_j])/\hat{h}(\theta^{**})$$

Idée 3 : Approximation de densités

On a : $[D|M_j] = ([D|\theta, M_j].[\theta|M_j])/[\theta|D, M_j]$
et ceci est vrai en particulier pour $\theta^* = E_{[\theta|D,M_j]}(\theta)$!

- Simuler $(\theta^1, \cdots \theta^T)$ sous la loi $[\theta|D, M_j]$ (MCMC !)
- Prendre un estimateur de $\theta^* \Rightarrow \theta^{**}=\sum \theta^t/T$
- Calcul de $([D|\theta^{**}, M_j].[\theta^{**}|M_j])$
- Estimation de la valeur de la densité h de $[\theta|D, M_j]$ en θ^{**} $\Rightarrow$ par ex, $\hat{h}(\theta^{**})$ estimateur a noyau gaussien (fonction density sous R) sur la base du MCMC

$$[D\hat{|}M_j] = ([D|\theta^{**}, M_j].[\theta^{**}|M_j])/\hat{h}(\theta^{**})$$

Idée 3 : Approximation de densités

On a : $[D|M_j] = ([D|\theta, M_j].[\theta|M_j])/[\theta|D, M_j]$
et ceci est vrai en particulier pour $\theta^* = E_{[\theta|D,M_j]}(\theta)$!

- Simuler $(\theta^1, \cdots \theta^T)$ sous la loi $[\theta|D, M_j]$ (MCMC !)
- Prendre un estimateur de $\theta^* \Rightarrow \theta^{**}$=$\sum \theta^t / T$
- Calcul de $([D|\theta^{**}, M_j].[\theta^{**}|M_j])$
- Estimation de la valeur de la densité h de $[\theta|D, M_j]$ en θ^{**} $\Rightarrow$ par ex, $\hat{h}(\theta^{**})$ estimateur a noyau gaussien (fonction density sous R) sur la base du MCMC

$$[D\hat{|}M_j] = ([D|\theta^{**}, M_j].[\theta^{**}|M_j])/\hat{h}(\theta^{**})$$

Idée 3 : Approximation de densités

On a : $[D|M_j] = ([D|\theta, M_j].[\theta|M_j])/[\theta|D, M_j]$
et ceci est vrai en particulier pour $\theta^* = E_{[\theta|D,M_j]}(\theta)$!

- Simuler $(\theta^1, \cdots \theta^T)$ sous la loi $[\theta|D, M_j]$ (MCMC !)
- Prendre un estimateur de $\theta^* \Rightarrow \theta^{**} = \sum \theta^t / T$
- Calcul de $([D|\theta^{**}, M_j].[\theta^{**}|M_j])$
- Estimation de la valeur de la densité h de $[\theta|D, M_j]$ en θ^{**} $\Rightarrow$ par ex, $\hat{h}(\theta^{**})$ estimateur a noyau gaussien (fonction density sous R) sur la base du MCMC

$$[D\hat{|}M_j] = ([D|\theta^{**}, M_j].[\theta^{**}|M_j])/\hat{h}(\theta^{**})$$

Idée 3 : Approximation de densités

On a : $[D|M_j] = ([D|\theta, M_j].[\theta|M_j])/[\theta|D, M_j]$
et ceci est vrai en particulier pour $\theta^* = E_{[\theta|D,M_j]}(\theta)$!

- Simuler $(\theta^1, \cdots \theta^T)$ sous la loi $[\theta|D, M_j]$ (MCMC !)
- Prendre un estimateur de $\theta^* \Rightarrow \theta^{**} = \sum \theta^t / T$
- Calcul de $([D|\theta^{**}, M_j].[\theta^{**}|M_j])$
- Estimation de la valeur de la densité h de $[\theta|D, M_j]$ en θ^{**}
  $\Rightarrow$ par ex, $\hat{h}(\theta^{**})$ estimateur a noyau gaussien (fonction density sous R) sur la base du MCMC

$$[D\hat{|}M_j] = ([D|\theta^{**}, M_j].[\theta^{**}|M_j])/\hat{h}(\theta^{**})$$

Idée 3 : Approximation de densités

On a : $[D|M_j] = ([D|\theta, M_j].[\theta|M_j])/[\theta|D, M_j]$
et ceci est vrai en particulier pour $\theta^* = E_{[\theta|D,M_j]}(\theta)$!

- Simuler $(\theta^1, \cdots \theta^T)$ sous la loi $[\theta|D, M_j]$ (MCMC !)
- Prendre un estimateur de $\theta^* \Rightarrow \theta^{**}=\sum \theta^t/T$
- Calcul de $([D|\theta^{**}, M_j].[\theta^{**}|M_j])$
- Estimation de la valeur de la densité h de $[\theta|D, M_j]$ en θ^{**} $\Rightarrow$ par ex, $\hat{h}(\theta^{**})$ estimateur a noyau gaussien (fonction density sous R) sur la base du MCMC

$$[D\hat{|}M_j] = ([D|\theta^{**}, M_j].[\theta^{**}|M_j])/\hat{h}(\theta^{**})$$

Exemple

Données

$X_1, ..., X_n$: une variable discrète positive
$Z_1, ..., Z_n$: une variable binaire

(par ex. : Z=1 pour groupe 1 , Z=2 pour groupe 2)

Modèle 1

$$X_i \sim Poisson(\lambda), \lambda \sim Gamma(\alpha, \beta)$$

Modèle 2

$$X_i \sim Poisson(\lambda_1 * I_{Zi=1} + \lambda_2 * I_{Zi=2})$$
$$\lambda_1 \sim Gamma(\alpha_1, \beta_1), \lambda_2 \sim Gamma(\alpha_2, \beta_2)$$

Remarques

- Sous les deux modèles $\Rightarrow$ les calculs sont explicites
- Cas 1 : Données simulées avec $n = 100$, $n_1 = n_2 = 50$ et $\lambda_1 = 2, \lambda_2 = 3$
- Cas 2 : Données simulées avec $n = 20$, $n_1 = n_2 = 10$ et $\lambda_1 = 2, \lambda_2 = 5$
- Deux lois a priori : $Gamma(1, 1)$ ou $Gamma(0.01, 0.01)$

Cas 1 où $\lambda_1 = 2$ et $\lambda_2 = 3$ (prior G(1,1))
$\Rightarrow Bf_{21} = 288$

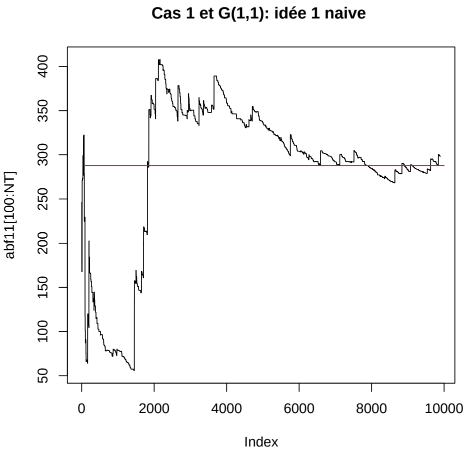


FIGURE: Estimations de BF21 avec Monte Carlo direct et prior G(1,1)

Cas 1 où $\lambda_1 = 2$ et $\lambda_2 = 3$ (prior G(1,1))
$\Rightarrow Bf_{21} = 288$

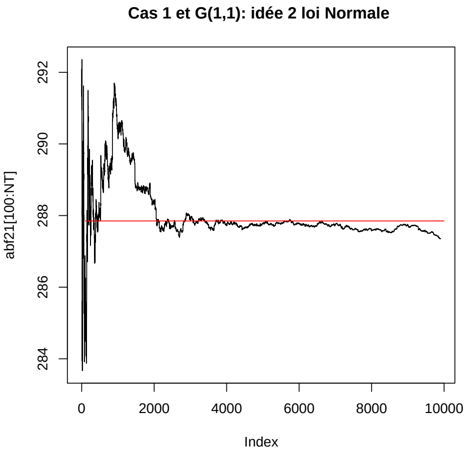


FIGURE: Estimations de BF21 avec Importance Sampling et prior G(1,1)

Cas 1 où $\lambda_1 = 2$ et $\lambda_2 = 3$ (prior G(1,1))
$\Rightarrow Bf_{21} = 288$

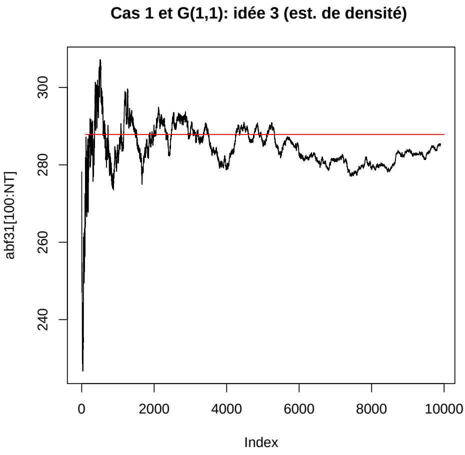


FIGURE: Estimations de BF21 avec estimation de densité et prior G(1,1)

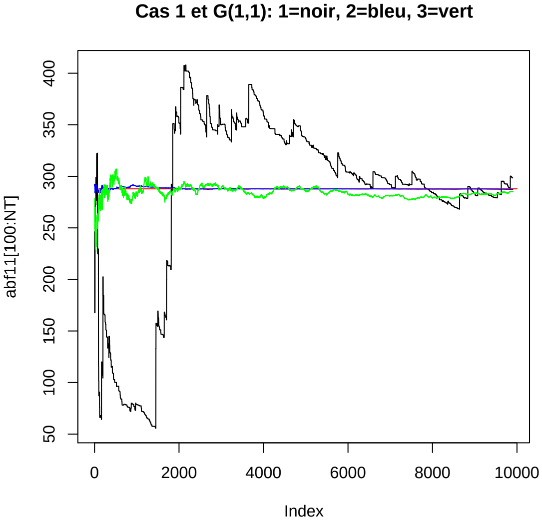


FIGURE: Graphe des estimations de BF21 avec prior G(1,1)

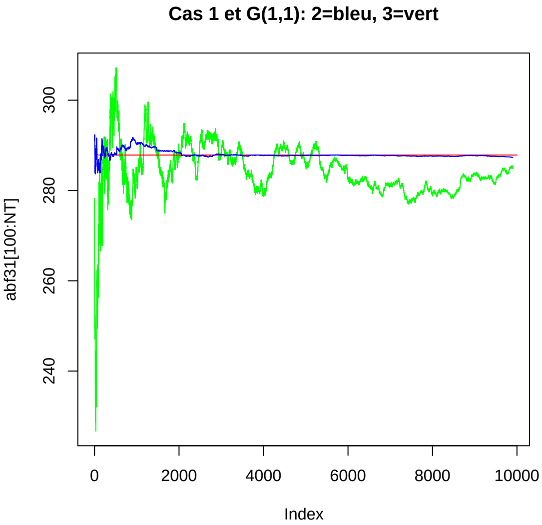
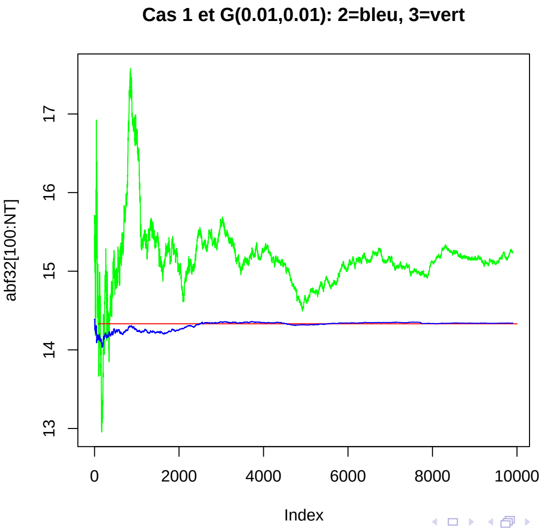
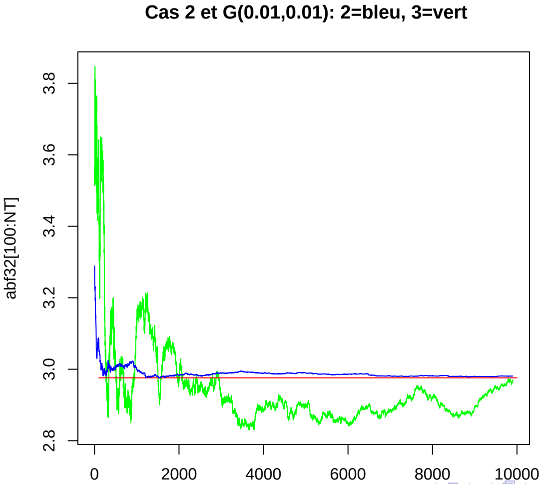


FIGURE: Graphe des estimations de BF21 avec prior G(1,1)

Cas 1 où $\lambda_1 = 2$ et $\lambda_2 = 3$ (prior vague)
$\Rightarrow Bf_{21} = 14$



Cas où les deux sous populations sont clairement séparées (avec prior vague) mais peu de données
$\Rightarrow BF_{21} \approx 3$, (P(Y|M1)=25% et P(Y|M2)=75%)



Exemple d'Eric Parent

Y : une variable binomiale

Modèle

$$Y \sim Bino(\theta, n) \quad \theta \sim Beta(a, b)$$

Simulation

a=6, b=4, y=93, n= 161

Les calculs sont explites et $log[y] = -4.290635$

Comment approcher $log[y]$?

$\Rightarrow$ Orienter sur le calcul de $[y]$ et non FB.

5 estimateurs de [y] proposés :

- y_1 = estimateur par approximation de densité (θ^{**}=mode a posteriori)
- y_2 = moyenne arithmétique (Monte Carlo direct)
- y_3 = importance sampling avec $g = prior$ (Raftery)
- y_4 = importance sampling avec $g = beta$
- y_5 = importance sampling avec $g = gaussienne$

Avec $m = 1000$ itérations,

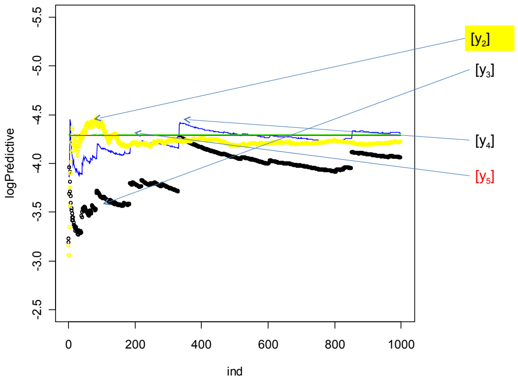


FIGURE: Graphe des estimations de [y] avec les 5 estimateurs

Avec $m = 1000000$ itérations,

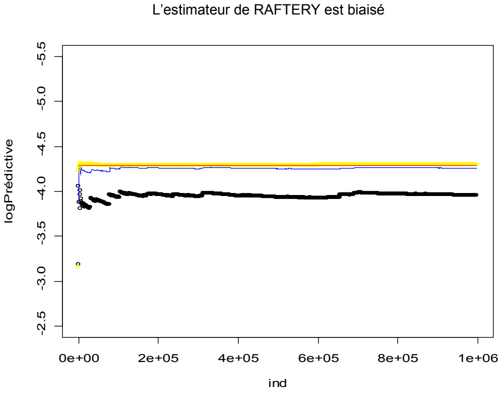


FIGURE: Graphe des estimations de [y] avec les 5 estimateurs

Performances en fonction du nombre d'itérations

$m = 100, 400, 1600, 6400, 25600, 102400, 409600, 817200, 1000000$

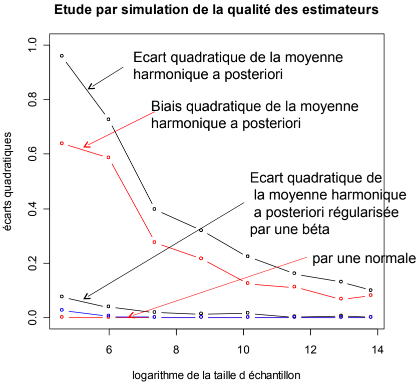


FIGURE: Performances des 5 estimateurs