

Bases de statistiques Bayésiennes

Point stats au CRPV

Marc Manceau

3 avril 2024

Motivations

Comprendre, critiquer, utiliser

Comprendre le jargon Bayésien :

- ▶ prior,
- ▶ posterior,
- ▶ vraisemblance,
- ▶ MCMC,
- ▶ prior conjugué,
- ▶ flat prior,
- ▶ ...

Lecture critique d'articles avec stats Bayésiennes :

- ▶ Comment critique-t-on de telles analyses ?
- ▶ Sur quels points des hypothèses / de l'analyse doit-on se montrer vigilant ?

Design de nos études :

- ▶ Dans quel cas de figure serait-il intéressant d'utiliser des stats bayésiennes vs. fréquentistes ?
- ▶ Comment réaliser les inférences le plus proprement possible ?

Motivations

Comprendre, critiquer, utiliser

Comprendre le jargon Bayésien :

- ▶ prior,
- ▶ posterior,
- ▶ vraisemblance,
- ▶ MCMC,
- ▶ prior conjugué,
- ▶ flat prior,
- ▶ ...

Lecture critique d'articles avec stats Bayésiennes :

- ▶ Comment critique-t-on de telles analyses ?
- ▶ Sur quels points des hypothèses / de l'analyse doit-on se montrer vigilant ?

Design de nos études :

- ▶ Dans quel cas de figure serait-il intéressant d'utiliser des stats bayésiennes vs. fréquentistes ?
- ▶ Comment réaliser les inférences le plus proprement possible ?

Motivations

Comprendre, critiquer, utiliser

Comprendre le jargon Bayésien :

- ▶ prior,
- ▶ posterior,
- ▶ vraisemblance,
- ▶ MCMC,
- ▶ prior conjugué,
- ▶ flat prior,
- ▶ ...

Lecture critique d'articles avec stats Bayésiennes :

- ▶ Comment critique-t-on de telles analyses ?
- ▶ Sur quels points des hypothèses / de l'analyse doit-on se montrer vigilant ?

Design de nos études :

- ▶ Dans quel cas de figure serait-il intéressant d'utiliser des stats bayésiennes vs. fréquentistes ?
- ▶ Comment réaliser les inférences le plus proprement possible ?

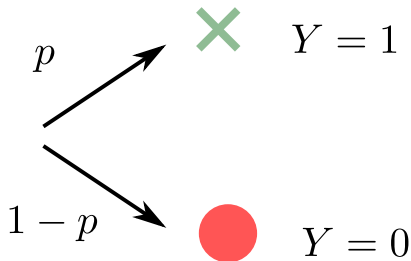
Motivations

Exemple filé

Exemple fictif un traitement utilisé pour lutter contre des encéphalées.

Outcome binaire : patient répondeur vs. non-répondeur.

Estimand la proportion de patients répondeurs dans une certaine population.



Motivations

Plan de présentation

Motivations

- Comprendre, critiquer, utiliser
- Exemple filé

Analyse fréquentiste

- Un modèle minimaliste : la vraisemblance
- Estimation et test

Analyse Bayésienne

- Un modèle plus complexe : prior & vraisemblance
- Déterminer la marginale des observations
- Déterminer le posterior
- Estimation et test

Outils associés

- Le MCMC, la cuisine du Bayésien
- Le prior conjugué, vers un modèle plus adapté/interprétable

Lecture critique

- Article 1
- Article 2

Conclusion

Analyse fréquentiste

Plan de présentation

Motivations

- Comprendre, critiquer, utiliser
- Exemple filé

Analyse fréquentiste

- Un modèle minimaliste : la vraisemblance
- Estimation et test

Analyse Bayésienne

- Un modèle plus complexe : prior & vraisemblance
- Déterminer la marginale des observations
- Déterminer le posterior
- Estimation et test

Outils associés

- Le MCMC, la cuisine du Bayésien
- Le prior conjugué, vers un modèle plus adapté/interprétable

Lecture critique

- Article 1
- Article 2

Conclusion

Analyse fréquentiste

Un modèle minimaliste : la vraisemblance

Modèle On tire n patients dans la même loi de Bernoulli de paramètre p .

Observation on observe le nombre de répondeurs : $N_r = 90$ avec 120 patients.

Vraisemblance la probabilité des observations étant donné un p fixé : $N_r \sim \mathcal{B}(n, p)$.

$$\mathbb{P}(N_r = 90 ; p) = \binom{120}{90} p^{90} (1 - p)^{30}$$



Analyse fréquentiste

Un modèle minimaliste : la vraisemblance

Modèle On tire n patients dans la même loi de Bernoulli de paramètre p .

Observation on observe le nombre de répondeurs : $N_r = 90$ avec 120 patients.

Vraisemblance la probabilité des observations étant donné un p fixé : $N_r \sim \mathcal{B}(n, p)$.

$$\mathbb{P}(N_r = 90 ; p) = \binom{120}{90} p^{90} (1 - p)^{30}$$



Analyse fréquentiste

Un modèle minimaliste : la vraisemblance

Modèle On tire n patients dans la même loi de Bernoulli de paramètre p .

Observation on observe le nombre de répondeurs : $N_r = 90$ avec 120 patients.

Vraisemblance la probabilité des observations étant donné un p fixé : $N_r \sim \mathcal{B}(n, p)$.

$$\mathbb{P}(N_r = 90 ; p) = \binom{120}{90} p^{90} (1 - p)^{30}$$



Analyse fréquentiste

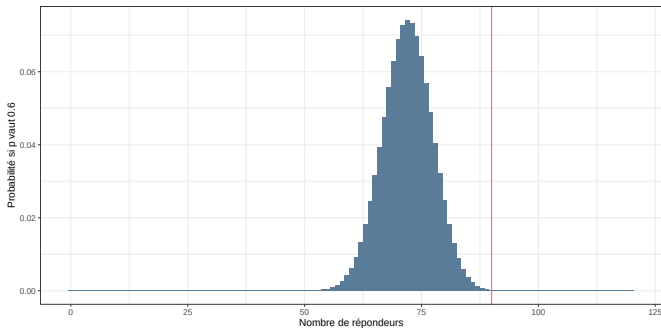
Un modèle minimaliste : la vraisemblance

Modèle On tire n patients dans la même loi de Bernoulli de paramètre p .

Observation on observe le nombre de répondeurs : $N_r = 90$ avec 120 patients.

Vraisemblance la probabilité des observations étant donné un p fixé : $N_r \sim \mathcal{B}(n, p)$.

$$\mathbb{P}(N_r = 90 ; p) = \binom{120}{90} p^{90} (1 - p)^{30}$$



Analyse fréquentiste

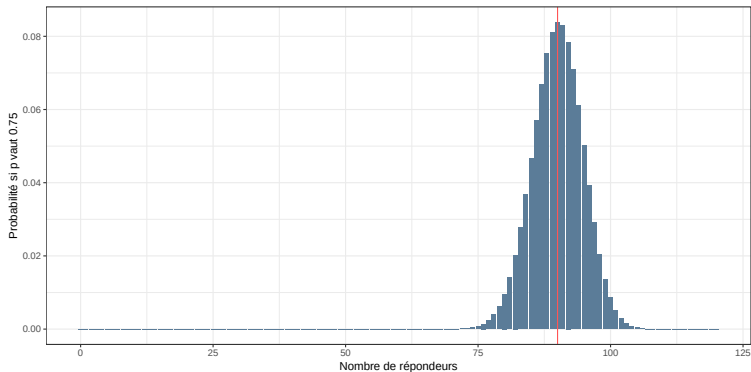
Estimation et test

Estimation ponctuelle par "maximum de vraisemblance" : $\hat{p} = \operatorname{argmax}_p \mathbb{P}(N_r ; p)$. Ici, $\hat{p} = N_r/n$.

Intervalle de confiance chercher $\hat{a}$ et $\hat{b}$ tels que $\mathbb{P}(p \in (\hat{a}, \hat{b})) \geq 95\%$

Test $H_0 : p \leq 0.6$ vs. $H_1 : p > 0.6$.

On décide sur la base de $\mathbb{P}(N_r \geq n_r ; p = 0.6)$.



Analyse fréquentiste

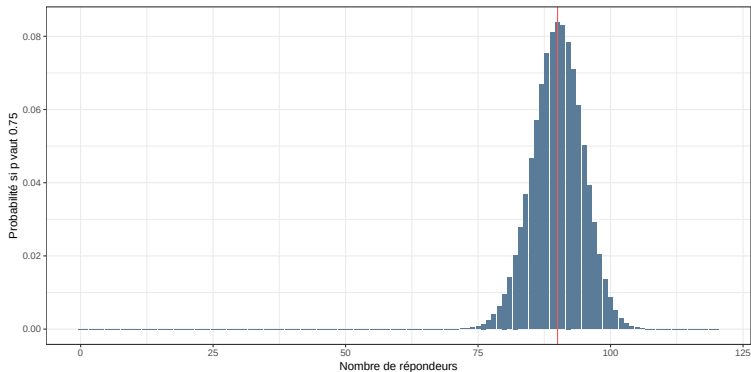
Estimation et test

Estimation ponctuelle par "maximum de vraisemblance" : $\hat{p} = \operatorname{argmax}_p \mathbb{P}(N_r ; p)$. Ici, $\hat{p} = N_r/n$.

Intervalle de confiance chercher $\hat{a}$ et $\hat{b}$ tels que $\mathbb{P}(p \in (\hat{a}, \hat{b})) \geq 95\%$

Test $H_0 : p \leq 0.6$ vs. $H_1 : p > 0.6$.

On décide sur la base de $\mathbb{P}(N_r \geq n_r ; p = 0.6)$.



Analyse fréquentiste

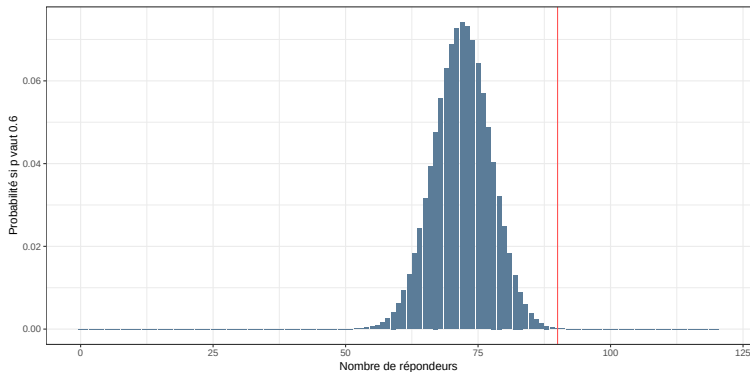
Estimation et test

Estimation ponctuelle par "maximum de vraisemblance" : $\hat{p} = \operatorname{argmax}_p \mathbb{P}(N_r ; p)$. Ici, $\hat{p} = N_r/n$.

Intervalle de confiance chercher $\hat{a}$ et $\hat{b}$ tels que $\mathbb{P}(p \in (\hat{a}, \hat{b})) \geq 95\%$

Test $H_0 : p \leq 0.6$ vs. $H_1 : p > 0.6$.

On décide sur la base de $\mathbb{P}(N_r \geq n_r ; p = 0.6)$.



Analyse Bayésienne

Plan de présentation

Motivations

- Comprendre, critiquer, utiliser
- Exemple filé

Analyse fréquentiste

- Un modèle minimaliste : la vraisemblance
- Estimation et test

Analyse Bayésienne

- Un modèle plus complexe : prior & vraisemblance
- Déterminer la marginale des observations
- Déterminer le posterior
- Estimation et test

Outils associés

- Le MCMC, la cuisine du Bayésien
- Le prior conjugué, vers un modèle plus adapté/interprétable

Lecture critique

- Article 1
- Article 2

Conclusion

Analyse Bayésienne

Un modèle plus complexe : prior & vraisemblance

Vraisemblance comme précédemment :

$$N_r | p \sim \mathcal{B}(n, p)$$

Prior la distribution de p , qui est supposé être aléatoire. Ex :

$$p \sim \mathcal{U}([0, 1])$$

Analyse Bayésienne

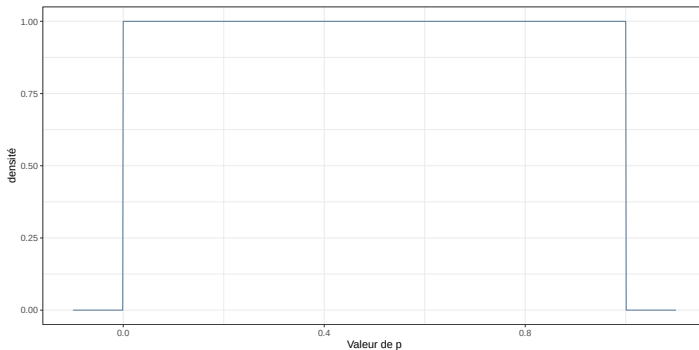
Un modèle plus complexe : prior & vraisemblance

Vraisemblance comme précédemment :

$$N_r|p \sim \mathcal{B}(n, p)$$

Prior la distribution de p , qui est supposé être aléatoire. Ex :

$$p \sim \mathcal{U}([0, 1])$$



Analyse Bayésienne

Déterminer la marginale des observations

Marginale des observations $\mathbb{P}(N_r) = \int_0^1 \mathbb{P}(N_r|p)\mathbb{P}(p \in dp)dp$

Pour l'explorer, on peut :

1. tirer une réalisation de p dans le prior,
2. tirer une réalisation de N_r dans sa loi conditionnelle à p (vraisemblance).

Analyse Bayésienne

Déterminer la marginale des observations

Marginale des observations $\mathbb{P}(N_r) = \int_0^1 \mathbb{P}(N_r|p)\mathbb{P}(p \in dp)dp$

Pour l'explorer, on peut :

1. tirer une réalisation de p dans le prior,
2. tirer une réalisation de N_r dans sa loi conditionnelle à p (vraisemblance).

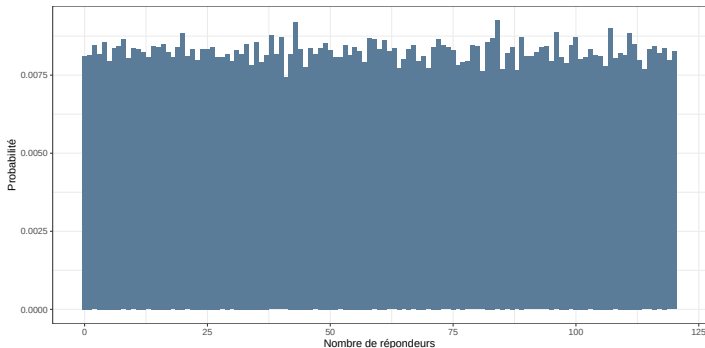
Analyse Bayésienne

Déterminer la marginale des observations

Marginale des observations $\mathbb{P}(N_r) = \int_0^1 \mathbb{P}(N_r|p)\mathbb{P}(p \in dp)dp$

Pour l'explorer, on peut :

1. tirer une réalisation de p dans le prior,
2. tirer une réalisation de N_r dans sa loi conditionnelle à p (vraisemblance).



Analyse Bayésienne

Déterminer le posterior

Observation le nombre de répondants : $N_r = 90$.

Posterior qui s'écrit à partir de la vraisemblance des données et du **prior**.

$$\mathbb{P}(p|N_r = 90)$$

Exemple avec notre prior uniforme, on trouve une loi Bêta : $\mathbb{P}(p|N_r = 90) \propto p^{90}(1 - p)^{30}$.

Analyse Bayésienne

Déterminer le posterior

Observation le nombre de répondants : $N_r = 90$.

Posterior qui s'écrit à partir de la **vraisemblance des données** et du **prior**.

$$\mathbb{P}(p|N_r = 90)$$

Exemple avec notre prior uniforme, on trouve une loi Bêta : $\mathbb{P}(p|N_r = 90) \propto p^{90}(1 - p)^{30}$.

Analyse Bayésienne

Déterminer le posterior

Observation le nombre de répondants : $N_r = 90$.

Posterior qui s'écrit à partir de la **vraisemblance des données** et du **prior**.

$$\mathbb{P}(p|N_r = 90) = \frac{\mathbb{P}(N_r = 90|p)\mathbb{P}(p)}{\mathbb{P}(N_r = 90)}$$

Exemple avec notre prior uniforme, on trouve une loi Bêta : $\mathbb{P}(p|N_r = 90) \propto p^{90}(1 - p)^{30}$.

Analyse Bayésienne

Déterminer le posterior

Observation le nombre de répondeurs : $N_r = 90$.

Posterior qui s'écrit à partir de la **vraisemblance des données** et du **prior**.

$$\mathbb{P}(p|N_r = 90) = \frac{\mathbb{P}(N_r = 90|p)\mathbb{P}(p)}{\mathbb{P}(N_r = 90)} \propto \mathbb{P}(N_r = 90|p)\mathbb{P}(p)$$

Exemple avec notre prior uniforme, on trouve une loi Bêta : $\mathbb{P}(p|N_r = 90) \propto p^{90}(1 - p)^{30}$.

Analyse Bayésienne

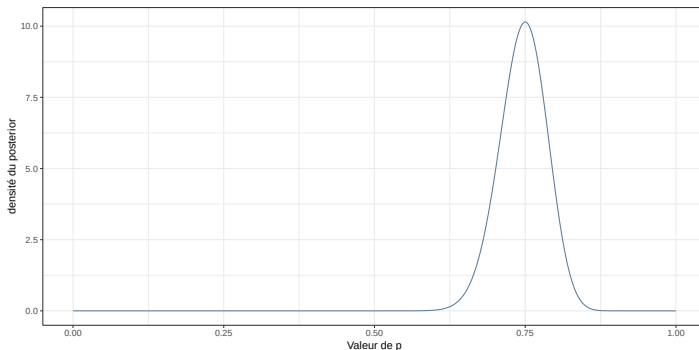
Déterminer le posterior

Observation le nombre de répondeurs : $N_r = 90$.

Posterior qui s'écrit à partir de la vraisemblance des données et du prior.

$$\mathbb{P}(p|N_r = 90) = \frac{\mathbb{P}(N_r = 90|p)\mathbb{P}(p)}{\mathbb{P}(N_r = 90)} \propto \mathbb{P}(N_r = 90|p)\mathbb{P}(p)$$

Exemple avec notre prior uniforme, on trouve une loi Bêta : $\mathbb{P}(p|N_r = 90) \propto p^{90}(1 - p)^{30}$.



Analyse Bayésienne

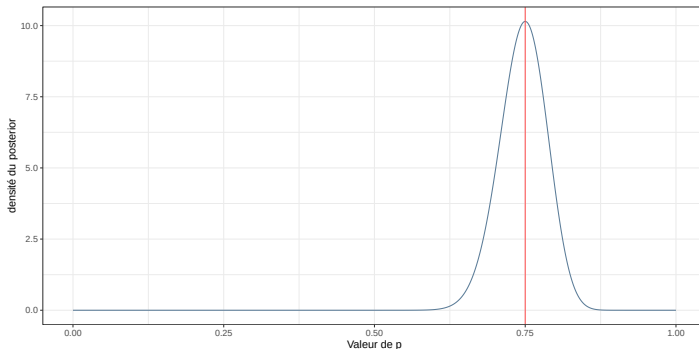
Estimation et test

Estimation ponctuelle avec le maximum a posteriori : $p_{MAP} = \operatorname{argmax}_p \mathbb{P}(p|N_r = 90)$.

Intervalle de crédibilité des quantiles a et b du posterior tels que :

$$\mathbb{P}(p \in (a, b) | N_r = 90) \geq 1 - \alpha$$

Test à partir de la probabilité a posteriori de différentes hypothèses (ex : $p \leq 0.7$).



Analyse Bayésienne

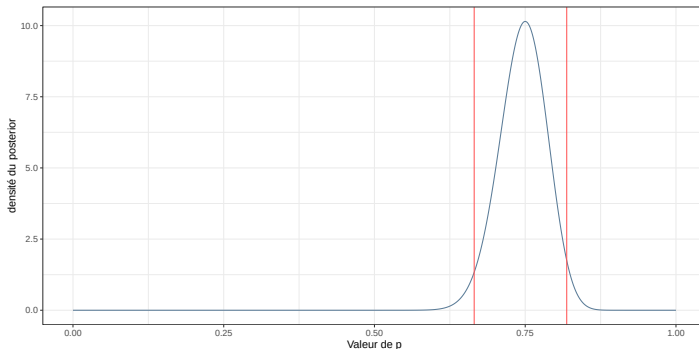
Estimation et test

Estimation ponctuelle avec le maximum a posteriori : $p_{MAP} = \operatorname{argmax}_p \mathbb{P}(p|N_r = 90)$.

Intervalle de crédibilité des quantiles a et b du posterior tels que :

$$\mathbb{P}(p \in (a, b) | N_r = 90) \geq 1 - \alpha$$

Test à partir de la probabilité a posteriori de différentes hypothèses (ex : $p \leq 0.7$).



Analyse Bayésienne

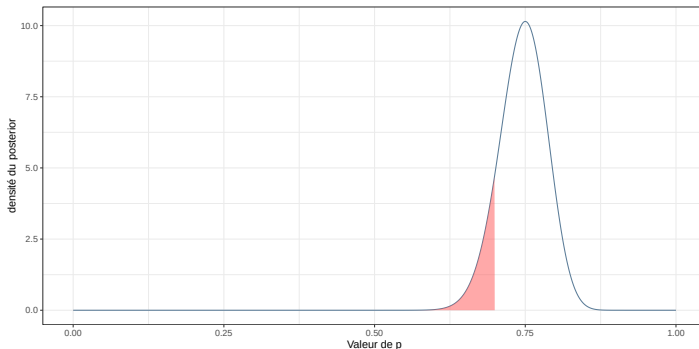
Estimation et test

Estimation ponctuelle avec le maximum a posteriori : $p_{MAP} = \operatorname{argmax}_p \mathbb{P}(p|N_r = 90)$.

Intervalle de crédibilité des quantiles a et b du posterior tels que :

$$\mathbb{P}(p \in (a, b) | N_r = 90) \geq 1 - \alpha$$

Test à partir de la probabilité a posteriori de différentes hypothèses (ex : $p \leq 0.7$).



Outils associés

Plan de présentation

Motivations

- Comprendre, critiquer, utiliser
- Exemple filé

Analyse fréquentiste

- Un modèle minimaliste : la vraisemblance
- Estimation et test

Analyse Bayésienne

- Un modèle plus complexe : prior & vraisemblance
- Déterminer la marginale des observations
- Déterminer le posterior
- Estimation et test

Outils associés

- Le MCMC, la cuisine du Bayésien
- Le prior conjugué, vers un modèle plus adapté/interprétable

Lecture critique

- Article 1
- Article 2

Conclusion

Outils associés

Le MCMC, la cuisine du Bayésien

Le prior est souvent ad-hoc :

- ▶ Des avis d'experts suggèrent que p serait centré sur 0.6, avec un écart-type de 0.3.
- ▶ On propose naïvement le prior suivant : $p \sim \mathcal{N}(0.6, 0.3)$.
- ▶ On décide de tronquer le prior pour qu'il ait pour support $[0, 1]$.

Outils associés

Le MCMC, la cuisine du Bayésien

Le prior est souvent ad-hoc :

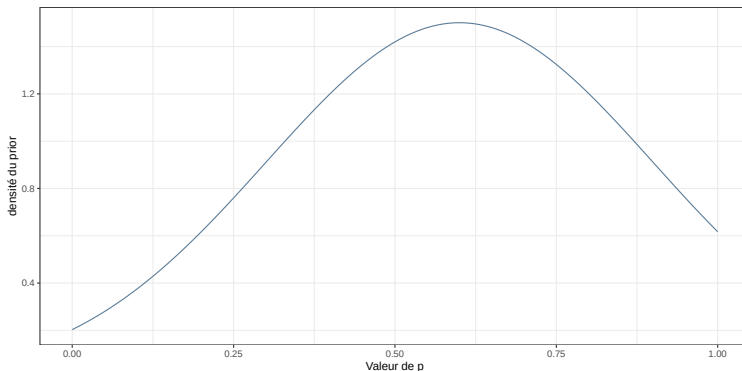
- ▶ Des avis d'experts suggèrent que p serait centré sur 0.6, avec un écart-type de 0.3.
- ▶ On propose naïvement le prior suivant : $p \sim \mathcal{N}(0.6, 0.3)$.
- ▶ On décide de tronquer le prior pour qu'il ait pour support $[0, 1]$.

Outils associés

Le MCMC, la cuisine du Bayésien

Le prior est souvent ad-hoc :

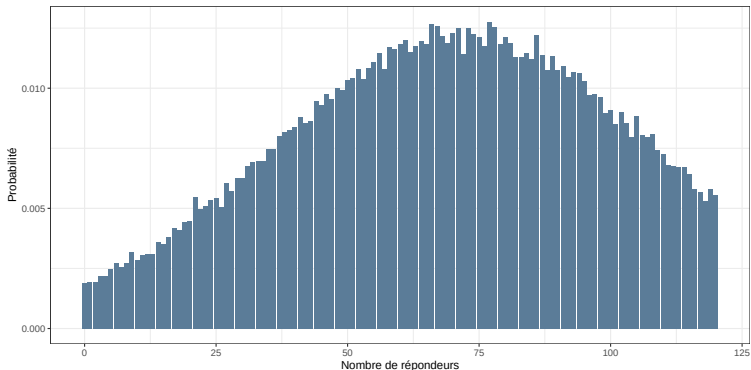
- ▶ Des avis d'experts suggèrent que p serait centré sur 0.6, avec un écart-type de 0.3.
- ▶ On propose naïvement le prior suivant : $p \sim \mathcal{N}(0.6, 0.3)$.
- ▶ On décide de tronquer le prior pour qu'il ait pour support $[0, 1]$.



Outils associés

Le MCMC, la cuisine du Bayésien

- On peut toujours échantillonner la marginale par simulations "naïves".

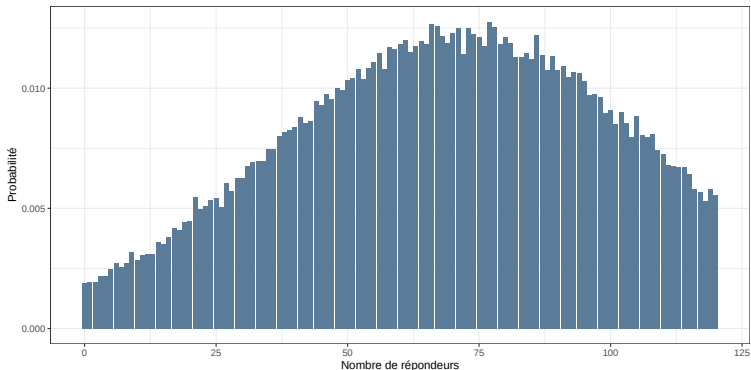


- Mais comment échantillonner le posterior ?

Outils associés

Le MCMC, la cuisine du Bayésien

- On peut toujours échantillonner la marginale par simulations "naïves".



- Mais comment échantillonner le posterior ?

Outils associés

Le MCMC, la cuisine du Bayésien

Chaîne de Markov une suite de variables aléatoires liées par $\mathbb{P}(X_{n+1}|X_n)$.

Objectif converger vers le posterior.

Algorithmes classiques Metropolis-Hastings, Gibbs sampling.



Algorithme de Metropolis-Hastings :

1. On fixe x_0 et une proposition de mouvement q .
 $\forall x, y, q(x, y)$ donne la proba de proposer l'état y sachant qu'on est dans l'état x .
2. Quand la chaîne est en x_i , on propose un état y en tirant dans la loi $q(x_i, \cdot)$.
3. Puis on calcule un genre d'"intérêt à aller explorer y ", avec ν le posterior :

$$r = \frac{\nu(y)q(y, x_i)}{\nu(x_i)q(x_i, y)}$$

4. Si $r \geq 1$, on accepte $x_{i+1} = y$.
Sinon $x_{i+1} = x_i$ avec probabilité r et $x_{i+1} = x_i$ sinon.

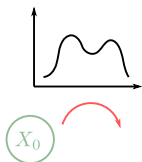
Outils associés

Le MCMC, la cuisine du Bayésien

Chaîne de Markov une suite de variables aléatoires liées par $\mathbb{P}(X_{n+1}|X_n)$.

Objectif converger vers le posterior.

Algorithmes classiques Metropolis-Hastings, Gibbs sampling.



Algorithme de Metropolis-Hastings :

1. On fixe x_0 et une proposition de mouvement q .
 $\forall x, y, q(x, y)$ donne la proba de proposer l'état y sachant qu'on est dans l'état x .
2. Quand la chaîne est en x_i , on propose un état y en tirant dans la loi $q(x_i, \cdot)$.
3. Puis on calcule un genre d'"intérêt à aller explorer y ", avec ν le posterior :

$$r = \frac{\nu(y)q(y, x_i)}{\nu(x_i)q(x_i, y)}$$

4. Si $r \geq 1$, on accepte $x_{i+1} = y$.
Sinon $x_{i+1} = x_i$ avec probabilité r et $x_{i+1} = x_i$ sinon.

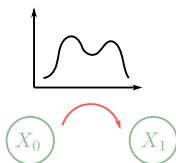
Outils associés

Le MCMC, la cuisine du Bayésien

Chaîne de Markov une suite de variables aléatoires liées par $\mathbb{P}(X_{n+1}|X_n)$.

Objectif converger vers le posterior.

Algorithmes classiques Metropolis-Hastings, Gibbs sampling.



Algorithme de Metropolis-Hastings :

1. On fixe x_0 et une proposition de mouvement q .
 $\forall x, y, q(x, y)$ donne la proba de proposer l'état y sachant qu'on est dans l'état x .
2. Quand la chaîne est en x_i , on propose un état y en tirant dans la loi $q(x_i, \cdot)$.
3. Puis on calcule un genre d'"intérêt à aller explorer y ", avec ν le posterior :

$$r = \frac{\nu(y)q(y, x_i)}{\nu(x_i)q(x_i, y)}$$

4. Si $r \geq 1$, on accepte $x_{i+1} = y$.
Sinon $x_{i+1} = y$ avec probabilité r et $x_{i+1} = x_i$ sinon.

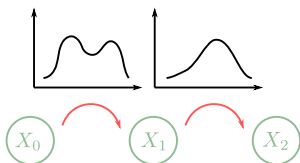
Outils associés

Le MCMC, la cuisine du Bayésien

Chaîne de Markov une suite de variables aléatoires liées par $\mathbb{P}(X_{n+1}|X_n)$.

Objectif converger vers le posterior.

Algorithmes classiques Metropolis-Hastings, Gibbs sampling.



Algorithme de Metropolis-Hastings :

1. On fixe x_0 et une proposition de mouvement q .
 $\forall x, y, q(x, y)$ donne la proba de proposer l'état y sachant qu'on est dans l'état x .
2. Quand la chaîne est en x_i , on propose un état y en tirant dans la loi $q(x_i, \cdot)$.
3. Puis on calcule un genre d'"intérêt à aller explorer y ", avec ν le posterior :

$$r = \frac{\nu(y)q(y, x_i)}{\nu(x_i)q(x_i, y)}$$

4. Si $r \geq 1$, on accepte $x_{i+1} = y$.
Sinon $x_{i+1} = y$ avec probabilité r et $x_{i+1} = x_i$ sinon.

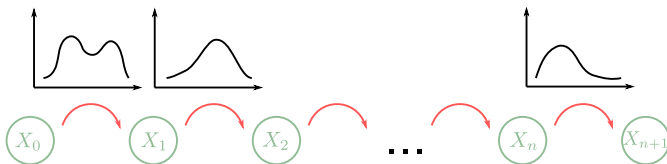
Outils associés

Le MCMC, la cuisine du Bayésien

Chaîne de Markov une suite de variables aléatoires liées par $\mathbb{P}(X_{n+1}|X_n)$.

Objectif converger vers le posterior.

Algorithmes classiques Metropolis-Hastings, Gibbs sampling.



Algorithme de Metropolis-Hastings :

1. On fixe x_0 et une proposition de mouvement q .
 $\forall x, y, q(x, y)$ donne la proba de proposer l'état y sachant qu'on est dans l'état x .
2. Quand la chaîne est en x_i , on propose un état y en tirant dans la loi $q(x_i, \cdot)$.
3. Puis on calcule un genre d'"intérêt à aller explorer y ", avec ν le posterior :

$$r = \frac{\nu(y)q(y, x_i)}{\nu(x_i)q(x_i, y)}$$

4. Si $r \geq 1$, on accepte $x_{i+1} = y$.
Sinon $x_{i+1} = x_i$ avec probabilité r et $x_{i+1} = x_i$ sinon.

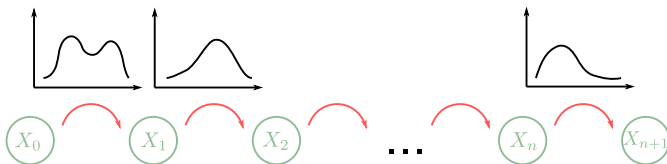
Outils associés

Le MCMC, la cuisine du Bayésien

Chaîne de Markov une suite de variables aléatoires liées par $\mathbb{P}(X_{n+1}|X_n)$.

Objectif converger vers le posterior.

Algorithmes classiques Metropolis-Hastings, Gibbs sampling.



Algorithme de Metropolis-Hastings :

1. On fixe x_0 et une proposition de mouvement q .
 $\forall x, y, q(x, y)$ donne la proba de proposer l'état y sachant qu'on est dans l'état x .
2. Quand la chaîne est en x_i , on propose un état y en tirant dans la loi $q(x_i, \cdot)$.
3. Puis on calcule un genre d'"intérêt à aller explorer y ", avec ν le posterior :

$$r = \frac{\nu(y)q(y, x_i)}{\nu(x_i)q(x_i, y)}$$

4. Si $r \geq 1$, on accepte $x_{i+1} = y$.
Sinon $x_{i+1} = x_i$ avec probabilité r et $x_{i+1} = x_i$ sinon.

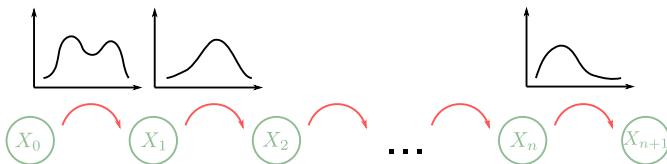
Outils associés

Le MCMC, la cuisine du Bayésien

Chaîne de Markov une suite de variables aléatoires liées par $\mathbb{P}(X_{n+1}|X_n)$.

Objectif converger vers le posterior.

Algorithmes classiques Metropolis-Hastings, Gibbs sampling.



Algorithme de Metropolis-Hastings :

1. On fixe x_0 et une proposition de mouvement q .
 $\forall x, y, q(x, y)$ donne la proba de proposer l'état y sachant qu'on est dans l'état x .
2. Quand la chaîne est en x_i , on propose un état y en tirant dans la loi $q(x_i, \cdot)$.
3. Puis on calcule un genre d'"intérêt à aller explorer y ", avec ν le posterior :

$$r = \frac{\nu(y)q(y, x_i)}{\nu(x_i)q(x_i, y)}$$

4. Si $r \geq 1$, on accepte $x_{i+1} = y$.
Sinon $x_{i+1} = x_i$ avec probabilité r et $x_{i+1} = x_i$ sinon.

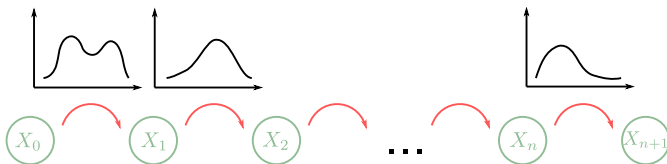
Outils associés

Le MCMC, la cuisine du Bayésien

Chaîne de Markov une suite de variables aléatoires liées par $\mathbb{P}(X_{n+1}|X_n)$.

Objectif converger vers le posterior.

Algorithmes classiques Metropolis-Hastings, Gibbs sampling.



Algorithme de Metropolis-Hastings :

1. On fixe x_0 et une proposition de mouvement q .
 $\forall x, y, q(x, y)$ donne la proba de proposer l'état y sachant qu'on est dans l'état x .
2. Quand la chaîne est en x_i , on propose un état y en tirant dans la loi $q(x_i, \cdot)$.
3. Puis on calcule un genre d'"intérêt à aller explorer y ", avec ν le posterior :

$$r = \frac{\nu(y)q(y, x_i)}{\nu(x_i)q(x_i, y)}$$

4. Si $r \geq 1$, on accepte $x_{i+1} = y$.
Sinon $x_{i+1} = x_i$ avec probabilité r et $x_{i+1} = x_i$ sinon.

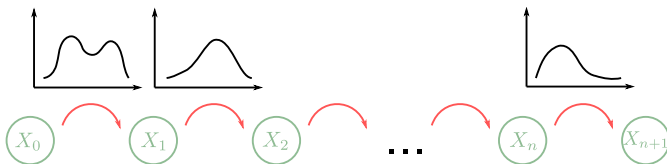
Outils associés

Le MCMC, la cuisine du Bayésien

Chaîne de Markov une suite de variables aléatoires liées par $\mathbb{P}(X_{n+1}|X_n)$.

Objectif converger vers le posterior.

Algorithmes classiques Metropolis-Hastings, Gibbs sampling.



Algorithme de Metropolis-Hastings :

1. On fixe x_0 et une proposition de mouvement q .
 $\forall x, y, q(x, y)$ donne la proba de proposer l'état y sachant qu'on est dans l'état x .
2. Quand la chaîne est en x_i , on propose un état y en tirant dans la loi $q(x_i, \cdot)$.
3. Puis on calcule un genre d'"intérêt à aller explorer y ", avec ν le posterior :

$$r = \frac{\nu(y)q(y, x_i)}{\nu(x_i)q(x_i, y)}$$

4. Si $r \geq 1$, on accepte $x_{i+1} = y$.
Sinon $x_{i+1} = x_i$ avec probabilité r et $x_{i+1} = x_i$ sinon.

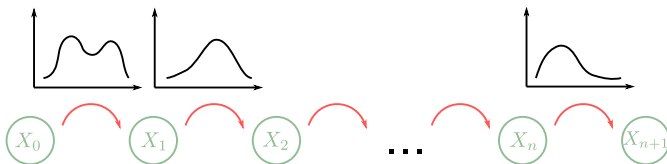
Outils associés

Le MCMC, la cuisine du Bayésien

Chaîne de Markov une suite de variables aléatoires liées par $\mathbb{P}(X_{n+1}|X_n)$.

Objectif converger vers le posterior.

Algorithmes classiques Metropolis-Hastings, Gibbs sampling.



Algorithme de Metropolis-Hastings :

1. On fixe x_0 et une proposition de mouvement q .
 $\forall x, y, q(x, y)$ donne la proba de proposer l'état y sachant qu'on est dans l'état x .
2. Quand la chaîne est en x_i , on propose un état y en tirant dans la loi $q(x_i, \cdot)$.
3. Puis on calcule un genre d'"intérêt à aller explorer y ", avec ν le posterior :

$$r = \frac{\nu(y)q(y, x_i)}{\nu(x_i)q(x_i, y)}$$

4. Si $r \geq 1$, on accepte $x_{i+1} = y$.
Sinon $x_{i+1} = x_i$ avec probabilité r et $x_{i+1} = x_i$ sinon.

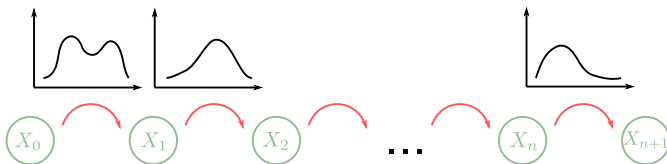
Outils associés

Le MCMC, la cuisine du Bayésien

Chaîne de Markov une suite de variables aléatoires liées par $\mathbb{P}(X_{n+1}|X_n)$.

Objectif converger vers le posterior.

Algorithmes classiques Metropolis-Hastings, Gibbs sampling.



Algorithme de Metropolis-Hastings :

1. On fixe x_0 et une proposition de mouvement q .
 $\forall x, y, q(x, y)$ donne la proba de proposer l'état y sachant qu'on est dans l'état x .
2. Quand la chaîne est en x_i , on propose un état y en tirant dans la loi $q(x_i, \cdot)$.
3. Puis on calcule un genre d'"intérêt à aller explorer y ", avec ν le posterior :

$$r = \frac{\nu(y)q(y, x_i)}{\nu(x_i)q(x_i, y)}$$

4. Si $r \geq 1$, on accepte $x_{i+1} = y$.
Sinon $x_{i+1} = x_i$ avec probabilité r et $x_{i+1} = x_i$ sinon.

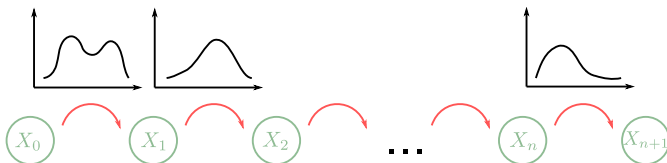
Outils associés

Le MCMC, la cuisine du Bayésien

Chaîne de Markov une suite de variables aléatoires liées par $\mathbb{P}(X_{n+1}|X_n)$.

Objectif converger vers le posterior.

Algorithmes classiques Metropolis-Hastings, Gibbs sampling.



Algorithme de Metropolis-Hastings :

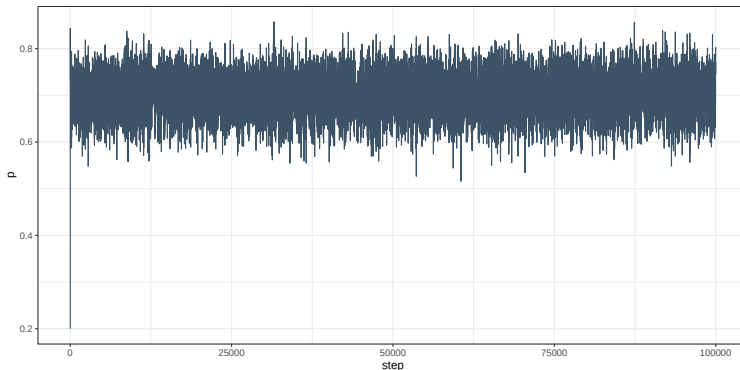
1. On fixe x_0 et une proposition de mouvement q .
 $\forall x, y, q(x, y)$ donne la proba de proposer l'état y sachant qu'on est dans l'état x .
2. Quand la chaîne est en x_i , on propose un état y en tirant dans la loi $q(x_i, \cdot)$.
3. Puis on calcule un genre d'"intérêt à aller explorer y ", avec ν le posterior :

$$r = \frac{\nu(y)q(y, x_i)}{\nu(x_i)q(x_i, y)}$$

4. Si $r \geq 1$, on accepte $x_{i+1} = y$.
 Sinon $x_{i+1} = x_i$ avec probabilité r et $x_{i+1} = x_i$ sinon.

Outils associés

Le MCMC, la cuisine du Bayésien



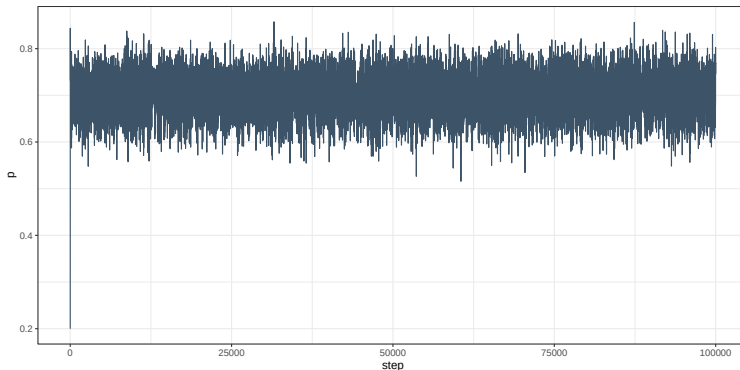
Dernières étapes :

burn-in on réduit la dépendance aux conditions initiales en supprimant les XX premières itérations.

échantillonnage on réduit l'auto-corrélation entre les valeurs en échantillonnant 1 valeur sur XX.

Outils associés

Le MCMC, la cuisine du Bayésien



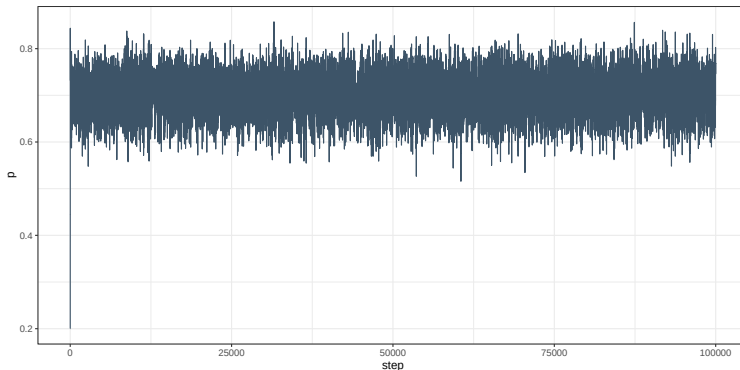
Dernières étapes :

burn-in on réduit la dépendance aux conditions initiales en supprimant les XX premières itérations.

échantillonnage on réduit l'auto-corrélation entre les valeurs en échantillonnant 1 valeur sur XX.

Outils associés

Le MCMC, la cuisine du Bayésien



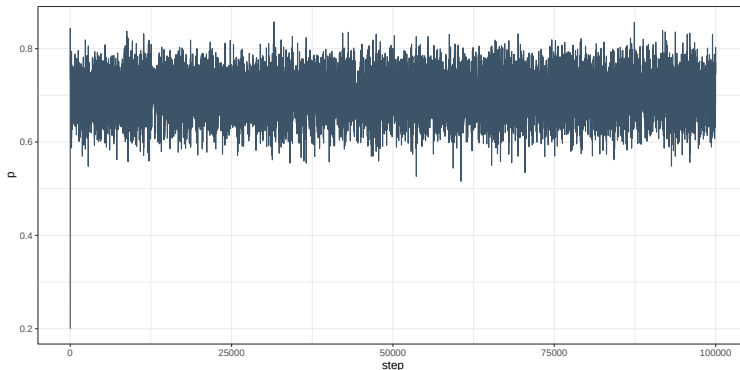
Dernières étapes :

burn-in on réduit la dépendance aux conditions initiales en supprimant les XX premières itérations.

échantillonnage on réduit l'auto-corrélation entre les valeurs en échantillonnant 1 valeur sur XX.

Outils associés

Le MCMC, la cuisine du Bayésien



Dernières étapes :

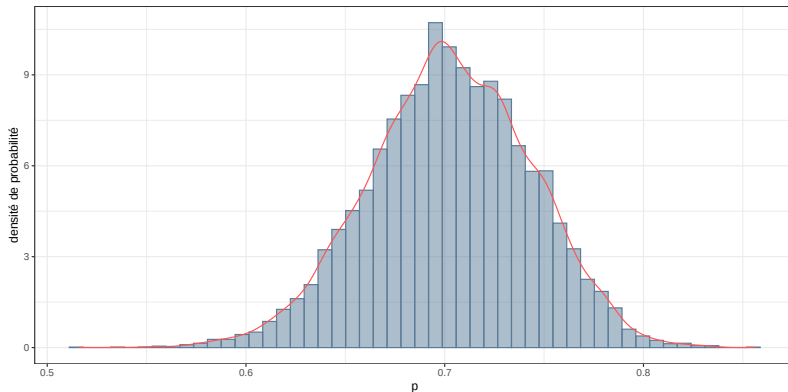
burn-in on réduit la dépendance aux conditions initiales en supprimant les XX premières itérations.

échantillonnage on réduit l'auto-corrélation entre les valeurs en échantillonnant 1 valeur sur XX.

Outils associés

Le MCMC, la cuisine du Bayésien

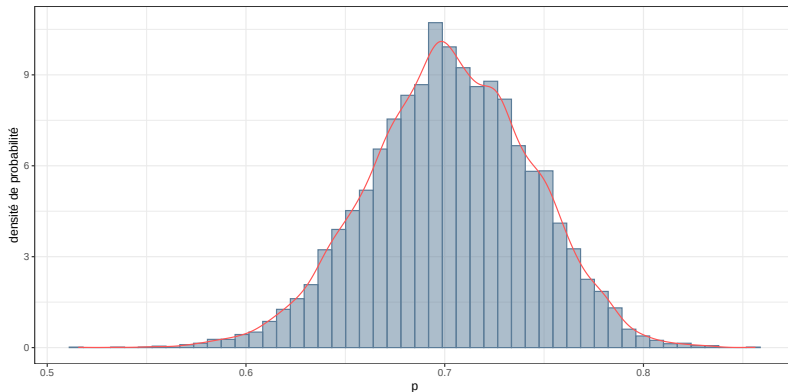
- ▶ On obtient notre estimation de posterior sans besoin de (trop) réfléchir.
- ▶ Au prix d'un certain nombre de choix ad-hoc et de calculs numériques.
- ▶ Quel est l'impact de ces choix ad-hoc sur le résultat ?



Outils associés

Le MCMC, la cuisine du Bayésien

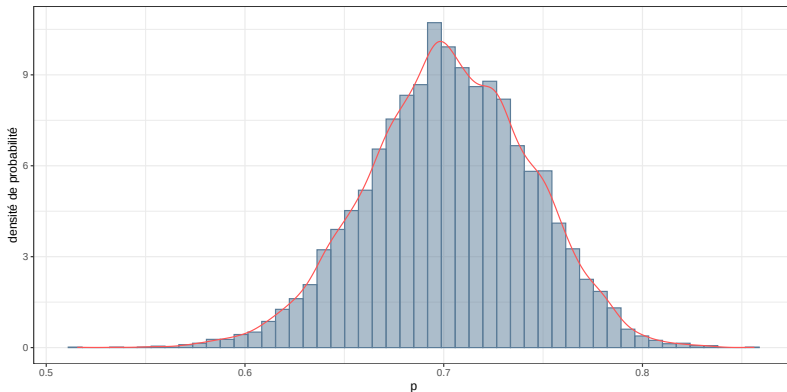
- ▶ On obtient notre estimation de posterior sans besoin de (trop) réfléchir.
- ▶ Au prix d'un certain nombre de choix ad-hoc et de calculs numériques.
- ▶ Quel est l'impact de ces choix ad-hoc sur le résultat ?



Outils associés

Le MCMC, la cuisine du Bayésien

- ▶ On obtient notre estimation de posterior sans besoin de (trop) réfléchir.
- ▶ Au prix d'un certain nombre de choix ad-hoc et de calculs numériques.
- ▶ Quel est l'impact de ces choix ad-hoc sur le résultat ?



Outils associés

Le prior conjugué, vers un modèle plus adapté/interprétable

Famille exponentielle Une classe de distributions contenant l'essentiel des distributions qu'on connaît.

Prior conjugué pour ces distributions, il existe un prior qui donne un posterior de même distribution.

Sur notre exemple la loi Bêta joue ce rôle :

$$\begin{aligned} &\text{si } N_r | p \sim \mathcal{B}(n, p) \\ &\text{et si } p \sim \text{Beta}(\alpha, \beta) \\ \implies &p | N_r \sim \text{Beta}(\alpha + N_r, \beta + n - N_r) \end{aligned}$$

Interprétabilité idéale de nos hyperparamètres en terme de nombre de réponders/non-réponders !

Outils associés

Le prior conjugué, vers un modèle plus adapté/interprétable

Famille exponentielle Une classe de distributions contenant l'essentiel des distributions qu'on connaît.

Prior conjugué pour ces distributions, il existe un prior qui donne un posterior de même distribution.

Sur notre exemple la loi Bêta joue ce rôle :

$$\begin{aligned} &\text{si } N_r | p \sim \mathcal{B}(n, p) \\ &\text{et si } p \sim \text{Beta}(\alpha, \beta) \\ \implies &p | N_r \sim \text{Beta}(\alpha + N_r, \beta + n - N_r) \end{aligned}$$

Interprétabilité idéale de nos hyperparamètres en terme de nombre de répondants/non-répondants !

Outils associés

Le prior conjugué, vers un modèle plus adapté/interprétable

Famille exponentielle Une classe de distributions contenant l'essentiel des distributions qu'on connaît.

Prior conjugué pour ces distributions, il existe un prior qui donne un posterior de même distribution.

Sur notre exemple la loi Bêta joue ce rôle :

$$\begin{aligned} \text{si } N_r | p &\sim \mathcal{B}(n, p) \\ \text{et si } p &\sim \text{Beta}(\alpha, \beta) \\ \implies p | N_r &\sim \text{Beta}(\alpha + N_r, \beta + n - N_r) \end{aligned}$$

Interprétabilité idéale de nos hyperparamètres en terme de nombre de répondeurs/non-répondeurs !

Outils associés

Le prior conjugué, vers un modèle plus adapté/interprétable

Famille exponentielle Une classe de distributions contenant l'essentiel des distributions qu'on connaît.

Prior conjugué pour ces distributions, il existe un prior qui donne un posterior de même distribution.

Sur notre exemple la loi Bêta joue ce rôle :

$$\begin{aligned} \text{si } N_r | p &\sim \mathcal{B}(n, p) \\ \text{et si } p &\sim \text{Beta}(\alpha, \beta) \\ \implies p | N_r &\sim \text{Beta}(\alpha + N_r, \beta + n - N_r) \end{aligned}$$

Interprétabilité idéale de nos hyperparamètres en terme de nombre de répondeurs/non-répondeurs !

Outils associés

Le prior conjugué, vers un modèle plus adapté/interprétable

- ▶ Une étude précédente s'est penchée sur le même problème.
- ▶ Ils avaient enregistré 50 répondeurs et 60 non-répondeurs.
- ▶ On peut partir sur un prior $p \sim \text{Beta}(\alpha = 50, \beta = 60)$.

Outils associés

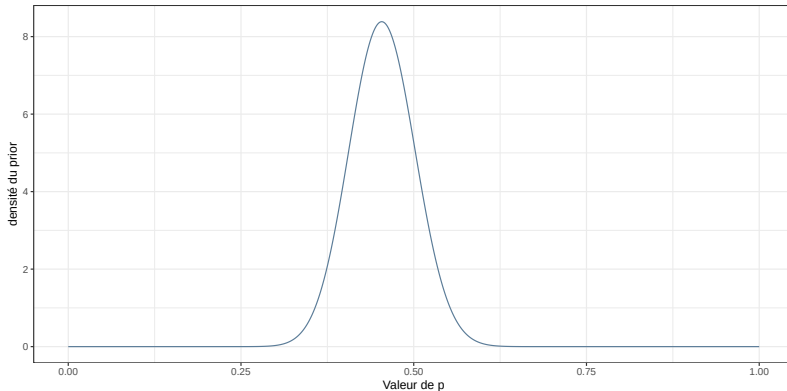
Le prior conjugué, vers un modèle plus adapté/interprétable

- ▶ Une étude précédente s'est penchée sur le même problème.
- ▶ Ils avaient enregistré 50 répondeurs et 60 non-répondeurs.
- ▶ On peut partir sur un prior $p \sim \text{Beta}(\alpha = 50, \beta = 60)$.

Outils associés

Le prior conjugué, vers un modèle plus adapté/interprétable

- ▶ Une étude précédente s'est penchée sur le même problème.
- ▶ Ils avaient enregistré 50 répondants et 60 non-répondants.
- ▶ On peut partir sur un prior $p \sim \text{Beta}(\alpha = 50, \beta = 60)$.

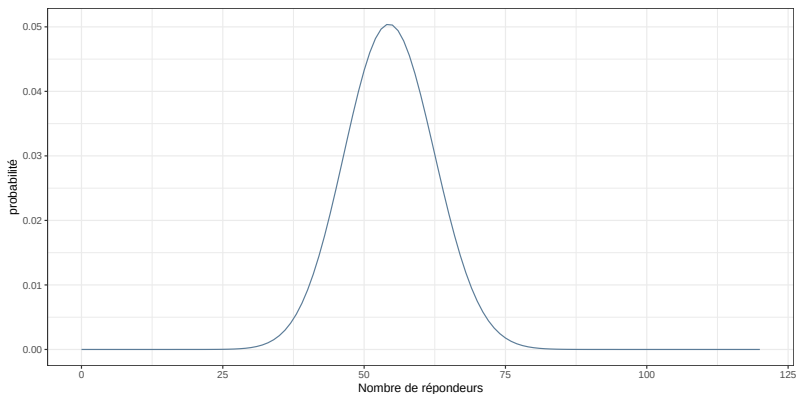


Outils associés

Le prior conjugué, vers un modèle plus adapté/interprétable

Marginale également connue avec des formules fermées (souvent compliquées).

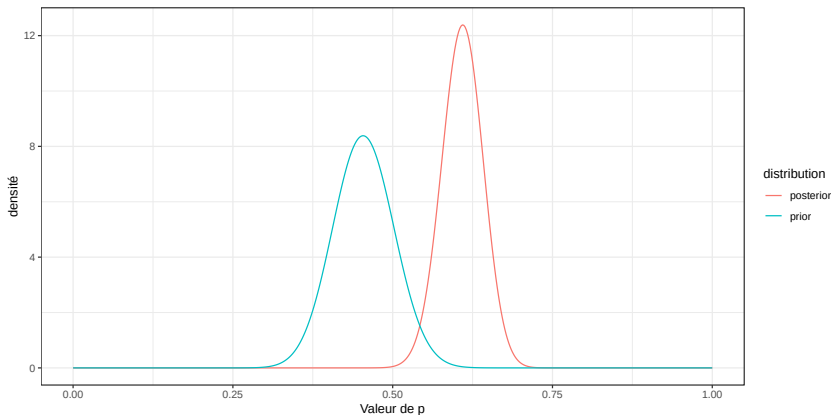
Sur notre exemple c'est une *Beta-Binomiale*.



Outils associés

Le prior conjugué, vers un modèle plus adapté/interprétable

Posterior facile à déterminer : $p|N_r \sim \text{Beta}(\alpha + N_r, \beta + n - N_r)$.



Lecture critique

Plan de présentation

Motivations

- Comprendre, critiquer, utiliser
- Exemple filé

Analyse fréquentiste

- Un modèle minimaliste : la vraisemblance
- Estimation et test

Analyse Bayésienne

- Un modèle plus complexe : prior & vraisemblance
- Déterminer la marginale des observations
- Déterminer le posterior
- Estimation et test

Outils associés

- Le MCMC, la cuisine du Bayésien
- Le prior conjugué, vers un modèle plus adapté/interprétable

Lecture critique

- Article 1
- Article 2

Conclusion

Lecture critique

Article 1

Research

JAMA Internal Medicine | [Original Investigation](#)

Effect of Tocilizumab vs Usual Care in Adults Hospitalized With COVID-19 and Moderate or Severe Pneumonia A Randomized Clinical Trial

Olivier Hermine, MD, PhD; Xavier Mariette, MD, PhD; Pierre-Louis Tharaux, MD, PhD; Matthieu Resche-Rigon, MD, PhD; Raphaël Porcher, PhD; Philippe Ravaud, MD, PhD; for the CORIMUNO-19 Collaborative Group

Patients COVID avec symptômes modérés ou sévères.

Bras TCZ et UC (usual care).

Outcome score d'une échelle de l'OMS (WHO-CPS) à J4, et survie sans besoin de ventilation à J14.

Recrutement 131 patients.

Lecture critique

Article 1 : les résultats

Outcome 1 : WHO-CPS score In the TCZ group, 12 patients had a WHO-CPS score greater than 5 at day 4 vs 19 in the UC group (median posterior absolute risk difference [ARD] -9.0% ; 90% credible interval [CrI], -21.0 to 3.1), with a posterior probability of negative ARD of 89.0% not achieving the 95% predefined efficacy threshold.

Outcome 2 : survival without ventilation At day 14, 12% (95% CI -28% to 4%) fewer patients needed noninvasive ventilation (NIV) or mechanical ventilation (MV) or died in the TCZ group than in the UC group (24% vs 36%, median posterior hazard ratio [HR] 0.58 ; 90% CrI, 0.33-1.00), with a posterior probability of HR less than 1 of 95.0%, achieving the predefined efficacy threshold.

Lecture critique

Article 1 : la cuisine Bayésienne

Interim analyses To maximize information from limited data generated while allowing for a rapid decision, we used a bayesian monitoring and analysis of the trial based on the coprimary outcomes. The sample size was set at 120, with interim analyses presented weekly to the DSMB and a provision to increase the sample size in case of promising but not conclusive results.

NSN We computed that the trial would have frequentist power of 97.2% to detect a decrease in event rate from 0.50 to 0.20, and 73.9% to detect a decrease in event rates from 0.50 to 0.30.

Priors For the day 4 outcome, we used a beta prior distribution with parameters 1 and 1 for the proportion in each arm (eFigure 1 in Supplement 3). For the day 14 outcome, we used a Gaussian prior distribution with a mean of 0 and variance of 106 for the log hazard ratio (HR).

Others Sensitivity analyses using a range of prior distribution were then conducted (eMethods, eFigure 2 in Supplement 3). The treatment effect was expressed in terms of absolute risk difference (ARD) for the day 4 outcome and HR for the day 14 outcome. Using Markov chain Monte Carlo methods, posterior probabilities of negative ARD and HR less than 1 were computed. According to the protocol, posterior probability greater than 0.99 at the interim analysis or greater than 0.95 at the final analysis indicated efficacy. We also computed posterior probability of ARD less than -5.5% and HR less than 0.85 (denoting moderate or greater effect). Because the decision rules are 1-sided, consistent credible intervals (Crls) would be theoretical. However, we chose to report 2-sided 90% Crls, which have the same upper bound as 1-sided 95% Crls.

Lecture critique

Article 1 : la cuisine Bayésienne

Interim analyses To maximize information from limited data generated while allowing for a rapid decision, we used a bayesian monitoring and analysis of the trial based on the coprimary outcomes. The sample size was set at 120, with interim analyses presented weekly to the DSMB and a provision to increase the sample size in case of promising but not conclusive results.

NSN We computed that the trial would have frequentist power of 97.2% to detect a decrease in event rate from 0.50 to 0.20, and 73.9% to detect a decrease in event rates from 0.50 to 0.30.

Priors For the day 4 outcome, we used a beta prior distribution with parameters 1 and 1 for the proportion in each arm (eFigure 1 in Supplement 3). For the day 14 outcome, we used a Gaussian prior distribution with a mean of 0 and variance of 106 for the log hazard ratio (HR).

Others Sensitivity analyses using a range of prior distribution were then conducted (eMethods, eFigure 2 in Supplement 3). The treatment effect was expressed in terms of absolute risk difference (ARD) for the day 4 outcome and HR for the day 14 outcome. Using Markov chain Monte Carlo methods, posterior probabilities of negative ARD and HR less than 1 were computed. According to the protocol, posterior probability greater than 0.99 at the interim analysis or greater than 0.95 at the final analysis indicated efficacy. We also computed posterior probability of ARD less than -5.5% and HR less than 0.85 (denoting moderate or greater effect). Because the decision rules are 1-sided, consistent credible intervals (CrIs) would be theoretical. However, we chose to report 2-sided 90% CrIs, which have the same upper bound as 1-sided 95% CrIs.

Points of concerns :

1. Design of interim analyses,
2. 5% one-sided Crl,
3. prior of HR leading to exactly 95% posterior proability ?

Lecture critique

Article 1 : la cuisine Bayésienne

Interim analyses To maximize information from limited data generated while allowing for a rapid decision, we used a bayesian monitoring and analysis of the trial based on the coprimary outcomes. The sample size was set at 120, with interim analyses presented weekly to the DSMB and a provision to increase the sample size in case of promising but not conclusive results.

NSN We computed that the trial would have frequentist power of 97.2% to detect a decrease in event rate from 0.50 to 0.20, and 73.9% to detect a decrease in event rates from 0.50 to 0.30.

Priors For the day 4 outcome, we used a beta prior distribution with parameters 1 and 1 for the proportion in each arm (eFigure 1 in Supplement 3). For the day 14 outcome, we used a Gaussian prior distribution with a mean of 0 and variance of 106 for the log hazard ratio (HR).

Others Sensitivity analyses using a range of prior distribution were then conducted (eMethods, eFigure 2 in Supplement 3). The treatment effect was expressed in terms of absolute risk difference (ARD) for the day 4 outcome and HR for the day 14 outcome. Using Markov chain Monte Carlo methods, posterior probabilities of negative ARD and HR less than 1 were computed. According to the protocol, posterior probability greater than 0.99 at the interim analysis or greater than 0.95 at the final analysis indicated efficacy. We also computed posterior probability of ARD less than -5.5% and HR less than 0.85 (denoting moderate or greater effect). Because the decision rules are 1-sided, consistent credible intervals (CrIs) would be theoretical. However, we chose to report 2-sided 90% CrIs, which have the same upper bound as 1-sided 95% CrIs.

Points of concerns :

1. Design of interim analyses,
2. 5% one-sided Crl,
3. prior of HR leading to exactly 95% posterior proability ?

Lecture critique

Article 1 : la cuisine Bayésienne

Interim analyses To maximize information from limited data generated while allowing for a rapid decision, we used a bayesian monitoring and analysis of the trial based on the coprimary outcomes. The sample size was set at 120, with interim analyses presented weekly to the DSMB and a provision to increase the sample size in case of promising but not conclusive results.

NSN We computed that the trial would have frequentist power of 97.2% to detect a decrease in event rate from 0.50 to 0.20, and 73.9% to detect a decrease in event rates from 0.50 to 0.30.

Priors For the day 4 outcome, we used a beta prior distribution with parameters 1 and 1 for the proportion in each arm (eFigure 1 in Supplement 3). For the day 14 outcome, we used a Gaussian prior distribution with a mean of 0 and variance of 106 for the log hazard ratio (HR).

Others Sensitivity analyses using a range of prior distribution were then conducted (eMethods, eFigure 2 in Supplement 3). The treatment effect was expressed in terms of absolute risk difference (ARD) for the day 4 outcome and HR for the day 14 outcome. Using Markov chain Monte Carlo methods, posterior probabilities of negative ARD and HR less than 1 were computed. According to the protocol, posterior probability greater than 0.99 at the interim analysis or greater than 0.95 at the final analysis indicated efficacy. We also computed posterior probability of ARD less than -5.5% and HR less than 0.85 (denoting moderate or greater effect). Because the decision rules are 1-sided, consistent credible intervals (CrIs) would be theoretical. However, we chose to report 2-sided 90% CrIs, which have the same upper bound as 1-sided 95% CrIs.

Points of concerns :

1. Design of interim analyses,
2. 5% one-sided Crl,
3. prior of HR leading to exactly 95% posterior proability ?

Lecture critique

Article 1 : la cuisine Bayésienne

Interim analyses To maximize information from limited data generated while allowing for a rapid decision, we used a bayesian monitoring and analysis of the trial based on the coprimary outcomes. The sample size was set at 120, with interim analyses presented weekly to the DSMB and a provision to increase the sample size in case of promising but not conclusive results.

NSN We computed that the trial would have frequentist power of 97.2% to detect a decrease in event rate from 0.50 to 0.20, and 73.9% to detect a decrease in event rates from 0.50 to 0.30.

Priors For the day 4 outcome, we used a beta prior distribution with parameters 1 and 1 for the proportion in each arm (eFigure 1 in Supplement 3). For the day 14 outcome, we used a Gaussian prior distribution with a mean of 0 and variance of 106 for the log hazard ratio (HR).

Others Sensitivity analyses using a range of prior distribution were then conducted (eMethods, eFigure 2 in Supplement 3). The treatment effect was expressed in terms of absolute risk difference (ARD) for the day 4 outcome and HR for the day 14 outcome. Using Markov chain Monte Carlo methods, posterior probabilities of negative ARD and HR less than 1 were computed. According to the protocol, posterior probability greater than 0.99 at the interim analysis or greater than 0.95 at the final analysis indicated efficacy. We also computed posterior probability of ARD less than -5.5% and HR less than 0.85 (denoting moderate or greater effect). Because the decision rules are 1-sided, consistent credible intervals (CrIs) would be theoretical. However, we chose to report 2-sided 90% CrIs, which have the same upper bound as 1-sided 95% CrIs.

Points of concerns :

1. Design of interim analyses,
2. 5% one-sided Crl,
3. prior of HR leading to exactly 95% posterior proability ?

Lecture critique

Article 1 : la cuisine Bayésienne

Interim analyses To maximize information from limited data generated while allowing for a rapid decision, we used a bayesian monitoring and analysis of the trial based on the coprimary outcomes. The sample size was set at 120, with interim analyses presented weekly to the DSMB and a provision to increase the sample size in case of promising but not conclusive results.

NSN We computed that the trial would have frequentist power of 97.2% to detect a decrease in event rate from 0.50 to 0.20, and 73.9% to detect a decrease in event rates from 0.50 to 0.30.

Priors For the day 4 outcome, we used a beta prior distribution with parameters 1 and 1 for the proportion in each arm (eFigure 1 in Supplement 3). For the day 14 outcome, we used a Gaussian prior distribution with a mean of 0 and variance of 106 for the log hazard ratio (HR).

Others Sensitivity analyses using a range of prior distribution were then conducted (eMethods, eFigure 2 in Supplement 3). The treatment effect was expressed in terms of absolute risk difference (ARD) for the day 4 outcome and HR for the day 14 outcome. Using Markov chain Monte Carlo methods, posterior probabilities of negative ARD and HR less than 1 were computed. According to the protocol, posterior probability greater than 0.99 at the interim analysis or greater than 0.95 at the final analysis indicated efficacy. We also computed posterior probability of ARD less than -5.5% and HR less than 0.85 (denoting moderate or greater effect). Because the decision rules are 1-sided, consistent credible intervals (CrIs) would be theoretical. However, we chose to report 2-sided 90% CrIs, which have the same upper bound as 1-sided 95% CrIs.

Points of concerns :

1. Design of interim analyses,
2. 5% one-sided CrI,
3. prior of HR leading to exactly 95% posterior probability?

Lecture critique

Article 2

Arthritis & Rheumatology

Vol. 72, No. 10, October 2020, pp 1621–1631

DOI 10.1002/art.41316

© 2020 Pfizer Inc. *Arthritis & Rheumatology* published by Wiley Periodicals LLC on behalf of American College of Rheumatology.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

AMERICAN COLLEGE
of RHEUMATOLOGY
Empowering Rheumatology Professionals

Efficacy and Safety of PF-06651600 (Ritlecitinib), a Novel JAK3/TEC Inhibitor, in Patients With Moderate-to-Severe Rheumatoid Arthritis and an Inadequate Response to Methotrexate

Michael F. Robinson,¹ Nemanja Damjanov,² Bojana Stamenkovic,³ Goran Radunovic,² Alan Kivitz,⁴ Lori Cox,⁵ Zorayr Manukyan,⁵ Christopher Banfield,⁵ Michael Saunders,⁵ Deepa Chandra,⁵ Michael S. Vincent,⁵ Jessica Mancuso,⁶ Elena Peeva,⁵ and Jean S. Beebe⁵

Patients avec une arthrite rhumatoïde.

Bras ritlecitinib vs. placebo.

Outcome Simplified Disease Activity Index (SDAI score) at W8.

Recrutement 70 patients.

Lecture critique

Article 2 : les méthodes

NSN The study planned to enroll 60 patients, expecting a 15% dropout rate to yield ~ 50 completers : 30 completers in the PF-06651600 group and 20 completers in the placebo group. The study had $>90\%$ power to detect a difference from placebo, assuming a 13-point difference in response between the PF-06651600 and placebo groups (the assumed SD was 14 points for both groups) and controlling the Type I error rate at 5%.

Model The primary efficacy end point analysis was conducted using a Bayesian analysis of covariance (ANCOVA) model with baseline SDAI score as a covariate.

Prior This model used an informative prior distribution for the placebo group (i.e., placebo data from recent tofacitinib trials with a similar patient population, study design, and efficacy/safety end points) after adjustment for baseline SDAI score and TNFi status from the present study. The prior model for placebo assumed that the mean SDAI response in the placebo group followed a normal distribution with a mean of -13.7 and variance of 25. Noninformative prior distribution was used for the mean SDAI response in the PF-06651600 group and for all other model parameters.

Choice of Bayesian analysis Utilizing historic placebo data from previous RA trials allowed for a decreased placebo sample size, thereby limiting the exposure of patients to placebo in this trial.

Sensitivity Multiple sensitivity analyses were considered for the prior mean of the placebo distribution. The main sensitivity analysis was conducted using noninformative prior distribution (i.e., without historic placebo information) for a normally distributed placebo response with a mean of 0 and variance of 1,000.

Lecture critique

Article 2 : les méthodes

NSN The study planned to enroll 60 patients, expecting a 15% dropout rate to yield ~ 50 completers : 30 completers in the PF-06651600 group and 20 completers in the placebo group. The study had $>90\%$ power to detect a difference from placebo, assuming a 13-point difference in response between the PF-06651600 and placebo groups (the assumed SD was 14 points for both groups) and **controlling the Type I error rate at 5%**.

Model The primary efficacy end point analysis was conducted using a Bayesian analysis of covariance (ANCOVA) model with baseline SDAI score as a covariate.

Prior This model used an informative prior distribution for the placebo group (i.e., placebo data from recent tofacitinib trials with a similar patient population, study design, and efficacy/safety end points) after adjustment for baseline SDAI score and TNFi status from the present study. The prior model for placebo assumed that the mean SDAI response in the placebo group followed a normal distribution with a mean of -13.7 and variance of 25. **Noninformative prior distribution** was used for the mean SDAI response in the PF-06651600 group and for all other model parameters.

Choice of Bayesian analysis Utilizing historic placebo data from previous RA trials **allowed for a decreased placebo sample size**, thereby limiting the exposure of patients to placebo in this trial.

Sensitivity Multiple sensitivity analyses were considered for the prior mean of the placebo distribution. The main sensitivity analysis was conducted using noninformative prior distribution (i.e., without historic placebo information) for a normally distributed placebo response with a mean of 0 and variance of 1,000.

Points of concerns :

1. NSN computed in frequentist setting, yet Bayesian analysis allowed for a decreased NSN ?
2. Priors could be better explained, including "non-informative" ones.
3. More patients included than what was planned.

Lecture critique

Article 2 : les méthodes

NSN The study planned to enroll 60 patients, expecting a 15% dropout rate to yield ~ 50 completers : 30 completers in the PF-06651600 group and 20 completers in the placebo group. The study had $>90\%$ power to detect a difference from placebo, assuming a 13-point difference in response between the PF-06651600 and placebo groups (the assumed SD was 14 points for both groups) and **controlling the Type I error rate at 5%**.

Model The primary efficacy end point analysis was conducted using a Bayesian analysis of covariance (ANCOVA) model with baseline SDAI score as a covariate.

Prior This model used an informative prior distribution for the placebo group (i.e., placebo data from recent tofacitinib trials with a similar patient population, study design, and efficacy/safety end points) after adjustment for baseline SDAI score and TNFi status from the present study. The prior model for placebo assumed that the mean SDAI response in the placebo group followed a normal distribution with a mean of -13.7 and variance of 25. **Noninformative prior distribution** was used for the mean SDAI response in the PF-06651600 group and for all other model parameters.

Choice of Bayesian analysis Utilizing historic placebo data from previous RA trials **allowed for a decreased placebo sample size**, thereby limiting the exposure of patients to placebo in this trial.

Sensitivity Multiple sensitivity analyses were considered for the prior mean of the placebo distribution. The main sensitivity analysis was conducted using noninformative prior distribution (i.e., without historic placebo information) for a normally distributed placebo response with a mean of 0 and variance of 1,000.

Points of concerns :

1. NSN computed in frequentist setting, yet Bayesian analysis allowed for a decreased NSN ?
2. Priors could be better explained, including "non-informative" ones.
3. More patients included than what was planned.

Lecture critique

Article 2 : les méthodes

NSN The study planned to enroll 60 patients, expecting a 15% dropout rate to yield ~ 50 completers : 30 completers in the PF-06651600 group and 20 completers in the placebo group. The study had $>90\%$ power to detect a difference from placebo, assuming a 13-point difference in response between the PF-06651600 and placebo groups (the assumed SD was 14 points for both groups) and **controlling the Type I error rate at 5%**.

Model The primary efficacy end point analysis was conducted using a Bayesian analysis of covariance (ANCOVA) model with baseline SDAI score as a covariate.

Prior This model used an informative prior distribution for the placebo group (i.e., placebo data from recent tofacitinib trials with a similar patient population, study design, and efficacy/safety end points) after adjustment for baseline SDAI score and TNFi status from the present study. The prior model for placebo assumed that the mean SDAI response in the placebo group followed a normal distribution with a mean of -13.7 and variance of 25. **Noninformative prior distribution** was used for the mean SDAI response in the PF-06651600 group and for all other model parameters.

Choice of Bayesian analysis Utilizing historic placebo data from previous RA trials **allowed for a decreased placebo sample size**, thereby limiting the exposure of patients to placebo in this trial.

Sensitivity Multiple sensitivity analyses were considered for the prior mean of the placebo distribution. The main sensitivity analysis was conducted using noninformative prior distribution (i.e., without historic placebo information) for a normally distributed placebo response with a mean of 0 and variance of 1,000.

Points of concerns :

1. NSN computed in frequentist setting, yet Bayesian analysis allowed for a decreased NSN ?
2. Priors could be better explained, including "non-informative" ones.
3. More patients included than what was planned.

Lecture critique

Article 2 : les méthodes

NSN The study planned to enroll 60 patients, expecting a 15% dropout rate to yield ~ 50 completers : 30 completers in the PF-06651600 group and 20 completers in the placebo group. The study had $>90\%$ power to detect a difference from placebo, assuming a 13-point difference in response between the PF-06651600 and placebo groups (the assumed SD was 14 points for both groups) and **controlling the Type I error rate at 5%**.

Model The primary efficacy end point analysis was conducted using a Bayesian analysis of covariance (ANCOVA) model with baseline SDAI score as a covariate.

Prior This model used an informative prior distribution for the placebo group (i.e., placebo data from recent tofacitinib trials with a similar patient population, study design, and efficacy/safety end points) after adjustment for baseline SDAI score and TNFi status from the present study. The prior model for placebo assumed that the mean SDAI response in the placebo group followed a normal distribution with a mean of -13.7 and variance of 25. **Noninformative prior distribution** was used for the mean SDAI response in the PF-06651600 group and for all other model parameters.

Choice of Bayesian analysis Utilizing historic placebo data from previous RA trials **allowed for a decreased placebo sample size**, thereby limiting the exposure of patients to placebo in this trial.

Sensitivity Multiple sensitivity analyses were considered for the prior mean of the placebo distribution. The main sensitivity analysis was conducted using noninformative prior distribution (i.e., without historic placebo information) for a normally distributed placebo response with a mean of 0 and variance of 1,000.

Points of concerns :

1. NSN computed in frequentist setting, yet Bayesian analysis allowed for a decreased NSN ?
2. Priors could be better explained, including "non-informative" ones.
3. More patients included than what was planned.

Lecture critique

Article 2 : les méthodes

NSN The study planned to enroll 60 patients, expecting a 15% dropout rate to yield ~ 50 completers : 30 completers in the PF-06651600 group and 20 completers in the placebo group. The study had $>90\%$ power to detect a difference from placebo, assuming a 13-point difference in response between the PF-06651600 and placebo groups (the assumed SD was 14 points for both groups) and **controlling the Type I error rate at 5%**.

Model The primary efficacy end point analysis was conducted using a Bayesian analysis of covariance (ANCOVA) model with baseline SDAI score as a covariate.

Prior This model used an informative prior distribution for the placebo group (i.e., placebo data from recent tofacitinib trials with a similar patient population, study design, and efficacy/safety end points) after adjustment for baseline SDAI score and TNFi status from the present study. The prior model for placebo assumed that the mean SDAI response in the placebo group followed a normal distribution with a mean of -13.7 and variance of 25. **Noninformative prior distribution** was used for the mean SDAI response in the PF-06651600 group and for all other model parameters.

Choice of Bayesian analysis Utilizing historic placebo data from previous RA trials **allowed for a decreased placebo sample size**, thereby limiting the exposure of patients to placebo in this trial.

Sensitivity Multiple sensitivity analyses were considered for the prior mean of the placebo distribution. The main sensitivity analysis was conducted using noninformative prior distribution (i.e., without historic placebo information) for a normally distributed placebo response with a mean of 0 and variance of 1,000.

Points of concerns :

1. NSN computed in frequentist setting, yet Bayesian analysis allowed for a decreased NSN ?
2. Priors could be better explained, including "non-informative" ones.
3. More patients included than what was planned.

Lecture critique

Article 2 : les résultats

Main analysis The primary efficacy end point, mean change from baseline in SDAI score at week 8, was greater with PF-06651600 (-26.1 [95% credible interval -29.7, -22.4]) than placebo (-16.8 [95% credible interval -20.9, -12.7]; $P < 0.001$) using a Bayesian informative prior distribution for placebo response.

Sensitivity A Bayesian ANCOVA with a noninformative prior distribution for placebo response confirmed the significant treatment effect based on the change from baseline in SDAI scores observed in the primary analysis ($P = 0.0023$). A parametric ANCOVA (using the week 8 end point) and a mixed-effects repeated-measures model (including all SDAI time points) were consistent with the primary analysis. Two modifications to the primary analysis population (excluding 5 patients after applying the inclusion/exclusion criteria and 4 patients who did not complete the week 8 visit) produced similar estimates of the significant treatment effect.

Conclusion

théorie	résultats plus simples demandant moins de bagage mathématique	théorie souvent plus ardue
interprétation	plus intuitive	moins intuitive
calcul numérique	besoin de plus de ressources numériques	plutôt plus rapide
besoin d'hypothèses	hypothèses fortes, dépendance aux priors	moins d'hypothèses nécessaires
	Bayésien	Fréquentiste

Quelques bonnes raisons d'utiliser un modèle Bayésien :

Prior informatif on souhaite intégrer des informations *a priori*.

Modèle très complexe pour lequel une inférence fréquentiste demanderait trop de matière grise.

Quelques bonnes pratiques en cas d'usage de stats Bayésiennes :

Justifier le(s) prior(s) et le(s) représenter graphiquement.

Choisir un prior adapté par exemple un prior conjugué.

Méthode d'inférence à documenter/fournir, avec d'éventuelles analyses de sensibilité.

Conclusion

théorie	résultats plus simples demandant moins de bagage mathématique	théorie souvent plus ardue
interprétation	plus intuitive	moins intuitive
calcul numérique	besoin de plus de ressources numériques	plutôt plus rapide
besoin d'hypothèses	hypothèses fortes, dépendance aux priors	moins d'hypothèses nécessaires
	Bayésien	Fréquentiste

Quelques bonnes raisons d'utiliser un modèle Bayésien :

Prior informatif on souhaite intégrer des informations *a priori*.

Modèle très complexe pour lequel une inférence fréquentiste demanderait trop de matière grise.

Quelques bonnes pratiques en cas d'usage de stats Bayésiennes :

Justifier le(s) prior(s) et le(s) représenter graphiquement.

Choisir un prior adapté par exemple un prior conjugué.

Méthode d'inférence à documenter/fournir, avec d'éventuelles analyses de sensibilité.

Conclusion

théorie	résultats plus simples demandant moins de bagage mathématique	théorie souvent plus ardue
interprétation	plus intuitive	moins intuitive
calcul numérique	besoin de plus de ressources numériques	plutôt plus rapide
besoin d'hypothèses	hypothèses fortes, dépendance aux priors	moins d'hypothèses nécessaires
	Bayésien	Fréquentiste

Quelques bonnes raisons d'utiliser un modèle Bayésien :

Prior informatif on souhaite intégrer des informations *a priori*.

Modèle très complexe pour lequel une inférence fréquentiste demanderait trop de matière grise.

Quelques bonnes pratiques en cas d'usage de stats Bayésiennes :

Justifier le(s) prior(s) et le(s) représenter graphiquement.

Choisir un prior adapté par exemple un prior conjugué.

Méthode d'inférence à documenter/fournir, avec d'éventuelles analyses de sensibilité.

Conclusion

théorie	résultats plus simples demandant moins de bagage mathématique	théorie souvent plus ardue
interprétation	plus intuitive	moins intuitive
calcul numérique	besoin de plus de ressources numériques	plutôt plus rapide
besoin d'hypothèses	hypothèses fortes, dépendance aux priors	moins d'hypothèses nécessaires
	Bayésien	Fréquentiste

Quelques bonnes raisons d'utiliser un modèle Bayésien :

Prior informatif on souhaite intégrer des informations *a priori*.

Modèle très complexe pour lequel une inférence fréquentiste demanderait trop de matière grise.

Quelques bonnes pratiques en cas d'usage de stats Bayésiennes :

Justifier le(s) prior(s) et le(s) représenter graphiquement.

Choisir un prior adapté par exemple un prior conjugué.

Méthode d'inférence à documenter/fournir, avec d'éventuelles analyses de sensibilité.

Conclusion

théorie	résultats plus simples demandant moins de bagage mathématique	théorie souvent plus ardue
interprétation	plus intuitive	moins intuitive
calcul numérique	besoin de plus de ressources numériques	plutôt plus rapide
besoin d'hypothèses	hypothèses fortes, dépendance aux priors	moins d'hypothèses nécessaires
	Bayésien	Fréquentiste

Quelques bonnes raisons d'utiliser un modèle Bayésien :

Prior informatif on souhaite intégrer des informations *a priori*.

Modèle très complexe pour lequel une inférence fréquentiste demanderait trop de matière grise.

Quelques bonnes pratiques en cas d'usage de stats Bayésiennes :

Justifier le(s) prior(s) et le(s) représenter graphiquement.

Choisir un prior adapté par exemple un prior conjugué.

Méthode d'inférence à documenter/fournir, avec d'éventuelles analyses de sensibilité.

Conclusion

théorie	résultats plus simples demandant moins de bagage mathématique	théorie souvent plus ardue
interprétation	plus intuitive	moins intuitive
calcul numérique	besoin de plus de ressources numériques	plutôt plus rapide
besoin d'hypothèses	hypothèses fortes, dépendance aux priors	moins d'hypothèses nécessaires
	Bayésien	Fréquentiste

Quelques bonnes raisons d'utiliser un modèle Bayésien :

Prior informatif on souhaite intégrer des informations *a priori*.

Modèle très complexe pour lequel une inférence fréquentiste demanderait trop de matière grise.

Quelques bonnes pratiques en cas d'usage de stats Bayésiennes :

Justifier le(s) prior(s) et le(s) représenter graphiquement.

Choisir un prior adapté par exemple un prior conjugué.

Méthode d'inférence à documenter/fournir, avec d'éventuelles analyses de sensibilité.

Conclusion

théorie	résultats plus simples demandant moins de bagage mathématique	théorie souvent plus ardue
interprétation	plus intuitive	moins intuitive
calcul numérique	besoin de plus de ressources numériques	plutôt plus rapide
besoin d'hypothèses	hypothèses fortes, dépendance aux priors	moins d'hypothèses nécessaires
	Bayésien	Fréquentiste

Quelques bonnes raisons d'utiliser un modèle Bayésien :

Prior informatif on souhaite intégrer des informations *a priori*.

Modèle très complexe pour lequel une inférence fréquentiste demanderait trop de matière grise.

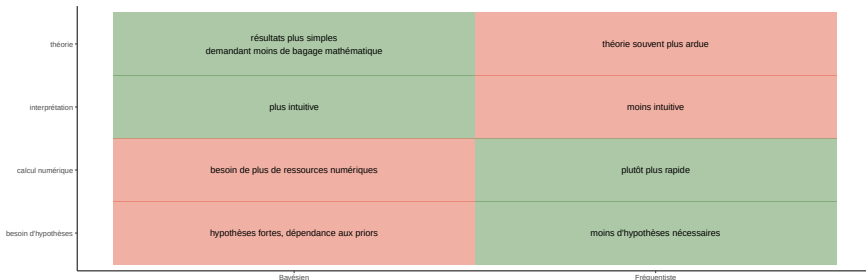
Quelques bonnes pratiques en cas d'usage de stats Bayésiennes :

Justifier le(s) prior(s) et le(s) représenter graphiquement.

Choisir un prior adapté par exemple un prior conjugué.

Méthode d'inférence à documenter/fournir, avec d'éventuelles analyses de sensibilité.

Conclusion



Quelques bonnes raisons d'utiliser un modèle Bayésien :

Prior informatif on souhaite intégrer des informations *a priori*.

Modèle très complexe pour lequel une inférence fréquentiste demanderait trop de matière grise.

Quelques bonnes pratiques en cas d'usage de stats Bayésiennes :

Justifier le(s) prior(s) et le(s) représenter graphiquement.

Choisir un prior adapté par exemple un prior conjugué.

Méthode d'inférence à documenter/fournir, avec d'éventuelles analyses de sensibilité.