Journées « modélisation en environnement »

Les modèles bayésiens hiérarchiques

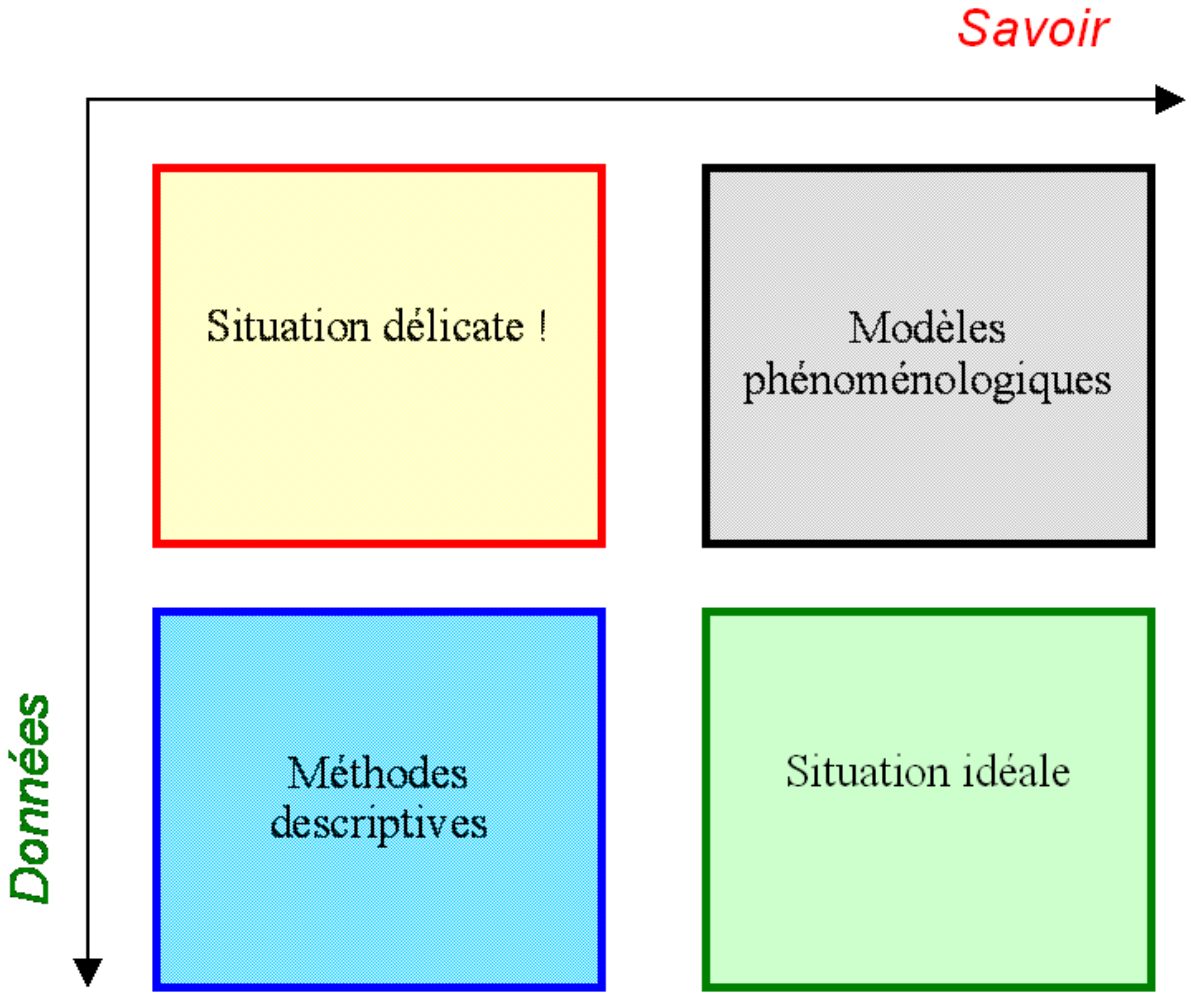
Jean-Jacques Boreux

ULg : Sciences et Gestion de l'environnement

22 février 2010

L'objet de la modélisation statistique

Interpréter des données à partir d'un savoir



Le travail du modélisateur

Mimer la nature, le plus souvent à des fins décisionnelles

- Le fabriquant de prêt-à-porter doit régler ses machines pour les jeans des garçonnets de 10 ans.

Le travail du modélisateur

Mimer la nature, le plus souvent à des fins décisionnelles

- Le fabriquant de prêt-à-porter doit régler ses machines pour les jeans des garçonnets de 10 ans.
  - Il a de l'expérience, i.e. son état de connaissance sur le problème est I.

Le travail du modélisateur

Mimer la nature, le plus souvent à des fins décisionnelles

- Le fabriquant de prêt-à-porter doit régler ses machines pour les jeans des garçonnets de 10 ans.
  - Il a de l'expérience, i.e. son état de connaissance sur le problème est I.
  - Il dispose aussi d'un échantillon (d'une mensuration caractéristique) généré par la "nature" :

$$D_n = \{y_i : i = 1, 2, \cdots, n\}$$

Le travail du modélisateur

Mimer la nature, le plus souvent à des fins décisionnelles

- Le fabriquant de prêt-à-porter doit régler ses machines pour les jeans des garçonnets de 10 ans.
  - Il a de l'expérience, i.e. son état de connaissance sur le problème est I.
  - Il dispose aussi d'un échantillon (d'une mensuration caractéristique) généré par la "nature" :

$$D_n = \{y_i : i = 1, 2, \cdots, n\}$$

- Le statisticien imagine que les données ont été tirées indépendamment les unes des autres dans une loi normale, localisée sur μ et de précision $\tau > 0$:

$$Y|\theta \sim dnorm\,(y|\theta)\,, \quad \theta = (\lambda, \tau)$$

Le travail du modélisateur

Mimer la nature, le plus souvent à des fins décisionnelles

- Le fabriquant de prêt-à-porter doit régler ses machines pour les jeans des garçonnets de 10 ans.
  - Il a de l'expérience, i.e. son état de connaissance sur le problème est I.
  - Il dispose aussi d'un échantillon (d'une mensuration caractéristique) généré par la "nature" :

$$D_n = \{y_i : i = 1, 2, \cdots, n\}$$

- Le statisticien imagine que les données ont été tirées indépendamment les unes des autres dans une loi normale, localisée sur μ et de précision $\tau > 0$:

$$Y|\theta \sim dnorm\,(y|\theta)\,, \quad \theta = (\lambda, \tau)$$

  - Pour le statisticien, la loi normale est un modèle d'échantillonnage qui "copie bien" la nature eu égard à la problématique en cours.

Le travail du modélisateur

Mimer la nature, le plus souvent à des fins décisionnelles

- Le fabriquant de prêt-à-porter doit régler ses machines pour les jeans des garçonnets de 10 ans.
  - Il a de l'expérience, i.e. son état de connaissance sur le problème est I.
  - Il dispose aussi d'un échantillon (d'une mensuration caractéristique) généré par la "nature" :

$$D_n = \{y_i : i = 1, 2, \cdots, n\}$$

- Le statisticien imagine que les données ont été tirées indépendamment les unes des autres dans une loi normale, localisée sur μ et de précision $\tau > 0$:

$$Y|\theta \sim dnorm\,(y|\theta)\,, \quad \theta = (\lambda, \tau)$$

  - Pour le statisticien, la loi normale est un modèle d'échantillonnage qui "copie bien" la nature eu égard à la problématique en cours.
  - Il doit maintenant identifier le membre de la famillenormale qui "colle" le mieux aux informations I et D_n.

L'inférence statistique

- Soit $D_n = (y_1, \cdots, y_n)$ un n échantillon *iid* selon une loi donnée :

$$Y|\theta \sim Loi\,(y|\theta)$$

L'inférence statistique

- Soit $D_n = (y_1, \cdots, y_n)$ un n échantillon *iid* selon une loi donnée :

$$Y|\theta \sim Loi\,(y|\theta)$$

- Le paramètre θ est inconnu, mais appartient à Θ : $\dim \Theta = d$ (finie).

L'inférence statistique

- Soit $D_n = (y_1, \cdots, y_n)$ un n échantillon *iid* selon une loi donnée :

$$Y|\theta \sim Loi\,(y|\theta)$$

- Le paramètre θ est inconnu, mais appartient à Θ : $\dim \Theta = d$ (finie).
- L'information apportée par D_n sur θ est condensée dans la fonction de vraisemblance :

$$[D_n|\theta] = \prod_{i=1}^{n} [y_i|\theta]$$

L'inférence statistique

- Soit $D_n = (y_1, \cdots, y_n)$ un n échantillon *iid* selon une loi donnée :

$$Y|\theta \sim Loi\,(y|\theta)$$

- Le paramètre θ est inconnu, mais appartient à Θ : $\dim \Theta = d$ (finie).
- L'information apportée par D_n sur θ est condensée dans la fonction de vraisemblance :

$$[D_n|\theta] = \prod_{i=1}^{n} [y_i|\theta]$$

  - $[y|\theta]$ est la fonction de densité de probabilité parente conditionnelle à θ.

L'inférence statistique

- Soit $D_n = (y_1, \cdots, y_n)$ un n échantillon *iid* selon une loi donnée :

$$Y|\theta \sim Loi\,(y|\theta)$$

- Le paramètre θ est inconnu, mais appartient à Θ : $\dim \Theta = d$ (finie).
- L'information apportée par D_n sur θ est condensée dans la fonction de vraisemblance :

$$[D_n|\theta] = \prod_{i=1}^{n} [y_i|\theta]$$

  - $[y|\theta]$ est la fonction de densité de probabilité parente conditionnelle à θ.

- L'apport majeur de l'approche bayésienne (par rapport à l'approche classique) est de substituer la distribution *a posteriori* à la vraisemblance :

$$[\theta|D_n, I] = \frac{[D_n|\theta]\,[\theta|I]}{\int_\Theta [D_n|\theta]\,[\theta|I]\,d\theta}$$

Le prior

- Avant de disposer de l'échantillon D_n, il est rare que l'on ne sache strictement rien du problème en main !

Le prior

- Avant de disposer de l'échantillon D_n, il est rare que l'on ne sache strictement rien du problème en main !
  - Exemple : les jours pluvieux à Marseille au mois de juillet.

Le prior

- Avant de disposer de l'échantillon D_n, il est rare que l'on ne sache strictement rien du problème en main !
  - Exemple : les jours pluvieux à Marseille au mois de juillet.
    - Les durées X_i séparant deux tops consécutifs sont *iid* selon une loi exponentielle :

$$[x_i|\lambda] = \lambda \exp\left(-\lambda x_i\right) \Rightarrow E\left(X_i|\lambda\right) = \lambda^{-1} > 0$$

Le prior

- Avant de disposer de l'échantillon D_n, il est rare que l'on ne sache strictement rien du problème en main !
  - Exemple : les jours pluvieux à Marseille au mois de juillet.
    - Les durées X_i séparant deux tops consécutifs sont *iid* selon une loi exponentielle :

$$[x_i|\lambda] = \lambda \exp\left(-\lambda x_i\right) \Rightarrow E\left(X_i|\lambda\right) = \lambda^{-1} > 0$$

    - Le spécialiste sait que λ ne peut pas prendre n'importe quelle valeur :

$$\Pr\left(\lambda^{-1} > 3.5\right) = 0.05, \quad \Pr\left(\lambda^{-1} < 1.5\right) = 0.2$$

Le prior

- Avant de disposer de l'échantillon D_n, il est rare que l'on ne sache strictement rien du problème en main !
  - Exemple : les jours pluvieux à Marseille au mois de juillet.
    - Les durées X_i séparant deux tops consécutifs sont *iid* selon une loi exponentielle :

$$[x_i|\lambda] = \lambda \exp\left(-\lambda x_i\right) \Rightarrow E\left(X_i|\lambda\right) = \lambda^{-1} > 0$$

    - Le spécialiste sait que λ ne peut pas prendre n'importe quelle valeur :

$$\Pr\left(\lambda^{-1} > 3.5\right) = 0.05, \quad \Pr\left(\lambda^{-1} < 1.5\right) = 0.2$$

    - Il peut alors représenter son expertise par un prior gamma :

$$[\lambda|a,b] = \frac{b^a}{\Gamma(a)}\lambda^{a-1}\exp\left(-b\lambda\right), \quad a \cong 9.8, \; b \cong 18.4$$

Le prior

- Avant de disposer de l'échantillon D_n, il est rare que l'on ne sache strictement rien du problème en main !
  - Exemple : les jours pluvieux à Marseille au mois de juillet.
    - Les durées X_i séparant deux tops consécutifs sont *iid* selon une loi exponentielle :

$$[x_i|\lambda] = \lambda \exp\left(-\lambda x_i\right) \Rightarrow E\left(X_i|\lambda\right) = \lambda^{-1} > 0$$

    - Le spécialiste sait que λ ne peut pas prendre n'importe quelle valeur :

$$\Pr\left(\lambda^{-1} > 3.5\right) = 0.05, \quad \Pr\left(\lambda^{-1} < 1.5\right) = 0.2$$

    - Il peut alors représenter son expertise par un prior gamma :

$$[\lambda|a, b] = \frac{b^a}{\Gamma(a)} \lambda^{a-1} \exp\left(-b\lambda\right), \quad a \cong 9.8, \ b \cong 18.4$$

    - La règle de Bayes donne :

$$\lambda|D_n, a, b \sim dgamma\left(\lambda|a+n, b+\Sigma x_i\right)$$

Le prior

- Avant de disposer de l'échantillon D_n, il est rare que l'on ne sache strictement rien du problème en main !
  - Exemple : les jours pluvieux à Marseille au mois de juillet.
    - Les durées X_i séparant deux tops consécutifs sont *iid* selon une loi exponentielle :

$$[x_i|\lambda] = \lambda \exp\left(-\lambda x_i\right) \Rightarrow E\left(X_i|\lambda\right) = \lambda^{-1} > 0$$

    - Le spécialiste sait que λ ne peut pas prendre n'importe quelle valeur :

$$\Pr\left(\lambda^{-1} > 3.5\right) = 0.05, \quad \Pr\left(\lambda^{-1} < 1.5\right) = 0.2$$

    - Il peut alors représenter son expertise par un prior gamma :

$$[\lambda|a, b] = \frac{b^a}{\Gamma(a)} \lambda^{a-1} \exp\left(-b\lambda\right), \quad a \cong 9.8, \ b \cong 18.4$$

    - La règle de Bayes donne :

$$\lambda|D_n, a, b \sim dgamma\left(\lambda|a+n, b+\Sigma x_i\right)$$

- Un prior vague permet d'encoder une expertise très réduite.

Du modèle statistique paramétrique à l'aide à la décision

- Exemple. On dispose du maximum annuel de la hauteur d'eau à l'échelle d'étiage sur la période $1980 - 2009$ (i.e. $n = 30$) et on a observé y dépassements du seuil critique c.

Du modèle statistique paramétrique à l'aide à la décision

- Exemple. On dispose du maximum annuel de la hauteur d'eau à l'échelle d'étiage sur la période $1980 - 2009$ (i.e. $n = 30$) et on a observé y dépassements du seuil critique c.
  - Le modèle d'échantillonnage est le modèle de Poisson ($\lambda = E\left(Y|n = 1\right)$) :

$$[y|\lambda, n] = \exp\left(-\lambda n\right) \frac{(\lambda n)^y}{y!}$$

Du modèle statistique paramétrique à l'aide à la décision

- Exemple. On dispose du maximum annuel de la hauteur d'eau à l'échelle d'étiage sur la période $1980 - 2009$ (i.e. $n = 30$) et on a observé y dépassements du seuil critique c.
  - Le modèle d'échantillonnage est le modèle de Poisson ($\lambda = E(Y|n = 1)$) :

$$[y|\lambda, n] = \exp(-\lambda n) \frac{(\lambda n)^y}{y!}$$

  - Le modèle de prior : le modèle gamma est conjugué au modèle de Poisson

$$[\lambda|a, b] = \frac{b^a}{\Gamma(a)} \lambda^{a-1} \exp(-b\lambda)$$

Du modèle statistique paramétrique à l'aide à la décision

- Exemple. On dispose du maximum annuel de la hauteur d'eau à l'échelle d'étiage sur la période $1980 - 2009$ (i.e. $n = 30$) et on a observé y dépassements du seuil critique c.
  - Le modèle d'échantillonnage est le modèle de Poisson ($\lambda = E(Y|n = 1)$) :

$$[y|\lambda, n] = \exp(-\lambda n) \frac{(\lambda n)^y}{y!}$$

  - Le modèle de prior : le modèle gamma est conjugué au modèle de Poisson

$$[\lambda|a, b] = \frac{b^a}{\Gamma(a)} \lambda^{a-1} \exp(-b\lambda)$$

  - Le posterior (il suffit d'appliquer la règle de Bayes) :

$$\lambda|n, y, a, b \sim dgamma(\lambda|y + a, n + b)$$

Du modèle statistique paramétrique à l'aide à la décision

- Exemple. On dispose du maximum annuel de la hauteur d'eau à l'échelle d'étiage sur la période $1980 - 2009$ (i.e. $n = 30$) et on a observé y dépassements du seuil critique c.
  - Le modèle d'échantillonnage est le modèle de Poisson $(\lambda = E(Y|n = 1))$:
$$[y|\lambda, n] = \exp(-\lambda n) \frac{(\lambda n)^y}{y!}$$
  - Le modèle de prior : le modèle gamma est conjugué au modèle de Poisson
$$[\lambda|a, b] = \frac{b^a}{\Gamma(a)} \lambda^{a-1} \exp(-b\lambda)$$
  - Le posterior (il suffit d'appliquer la règle de Bayes) :
$$\lambda|n, y, a, b \sim dgamma(\lambda|y + a, n + b)$$
  - Tout questionnement ultérieur sera basé sur la distribution *a posteriori* du paramètre de Poisson !

Du modèle statistique paramétrique à l'aide à la décision

- Exemple. On dispose du maximum annuel de la hauteur d'eau à l'échelle d'étiage sur la période $1980 - 2009$ (i.e. $n = 30$) et on a observé y dépassements du seuil critique c.
  - Le modèle d'échantillonnage est le modèle de Poisson $(\lambda = E(Y|n = 1))$:
$$[y|\lambda, n] = \exp(-\lambda n) \frac{(\lambda n)^y}{y!}$$
  - Le modèle de prior : le modèle gamma est conjugué au modèle de Poisson
$$[\lambda|a, b] = \frac{b^a}{\Gamma(a)} \lambda^{a-1} \exp(-b\lambda)$$
  - Le posterior (il suffit d'appliquer la règle de Bayes) :
$$\lambda|n, y, a, b \sim dgamma(\lambda|y + a, n + b)$$
  - Tout questionnement ultérieur sera basé sur la distribution *a posteriori* du paramètre de Poisson !
- $\Rightarrow$ distribution prédictive *a posteriori.*

La distribution prédictive a posteriori

- Exemple (continué). Quelles sont les chances que l'on oberve au moins un dépassement dans les h prochaines années ?

La distribution prédictive a posteriori

- Exemple (continué). Quelles sont les chances que l'on oberve au moins un dépassement dans les h prochaines années ?
  - La distribution prédictive *a posteriori* est la loi binomiale négative ($f \equiv futur, p \equiv passé$)

$$
\begin{aligned}
[Y = y_f | h, Y = y_p, n] &= \int_\Theta [Y = y_f, \lambda | h, Y = y_p, n]\, d\lambda \\
&= \int_\Theta \underbrace{[Y = y_f | \lambda, h]}_{Poisson} \times \underbrace{[\lambda | Y = y_p, n]}_{Gamma} d\lambda \\
&= \frac{h^{y_f}}{y_f!} \times \frac{(n+b)^{y_p+a}}{\Gamma(y_p+a)} \frac{\Gamma(y_f+y_p+a)}{(h+n+b)^{y_f+y_p+a}}
\end{aligned}
$$

La distribution prédictive a posteriori

- Exemple (continué). Quelles sont les chances que l'on oberve au moins un dépassement dans les h prochaines années ?
  - La distribution prédictive *a posteriori* est la loi binomiale négative ($f \equiv futur, p \equiv passé$)

$$
\begin{aligned}
[Y = y_f | h, Y = y_p, n] &= \int_\Theta [Y = y_f, \lambda | h, Y = y_p, n]\, d\lambda \\
&= \int_\Theta \underbrace{[Y = y_f | \lambda, h]}_{Poisson} \times \underbrace{[\lambda | Y = y_p, n]}_{Gamma} d\lambda \\
&= \frac{h^{y_f}}{y_f!} \times \frac{(n+b)^{y_p+a}}{\Gamma(y_p+a)} \frac{\Gamma(y_f+y_p+a)}{(h+n+b)^{y_f+y_p+a}}
\end{aligned}
$$

  - Finalement

$$
[Y > 0 | h, Y = y_p, n] = 1 - \frac{1}{\left(\frac{h}{n+b} + 1\right)^{y_p+a}}
$$

La distribution prédictive a posteriori

- Exemple (continué). Quelles sont les chances que l'on oberve au moins un dépassement dans les h prochaines années ?
  - La distribution prédictive *a posteriori* est la loi binomiale négative ($f \equiv futur, p \equiv passé$)

$$
\begin{aligned}
[Y = y_f | h, Y = y_p, n] &= \int_\Theta [Y = y_f, \lambda | h, Y = y_p, n]\, d\lambda \\
&= \int_\Theta \underbrace{[Y = y_f | \lambda, h]}_{Poisson} \times \underbrace{[\lambda | Y = y_p, n]}_{Gamma} d\lambda \\
&= \frac{h^{y_f}}{y_f!} \times \frac{(n+b)^{y_p+a}}{\Gamma(y_p+a)} \frac{\Gamma(y_f+y_p+a)}{(h+n+b)^{y_f+y_p+a}}
\end{aligned}
$$

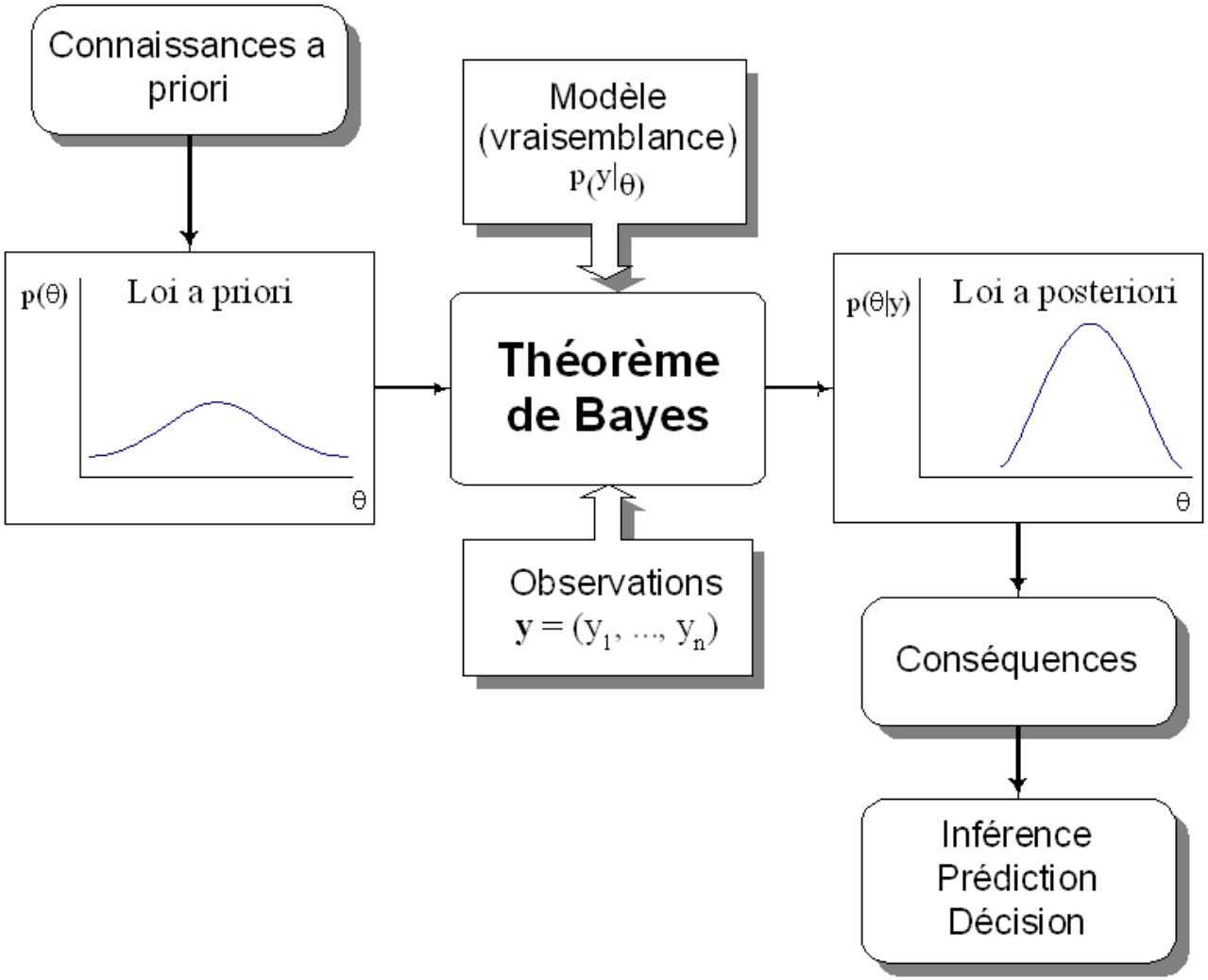
  - Finalement

$$
[Y > 0 | h, Y = y_p, n] = 1 - \frac{1}{\left(\frac{h}{n+b} + 1\right)^{y_p+a}}
$$

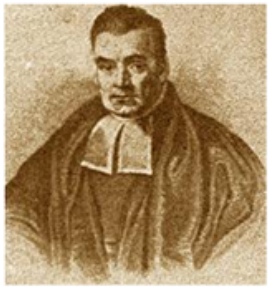
- La distribution prédictive *a posteriori* n'a pas d'équivalent classique !

Le paradigme bayésien

Résumé



Évolution de la statistique bayésienne



Théorème de Bayes

$$p(H|D) = \frac{p(D|H)\,p(H)}{p(D)}$$

1763

$$p(\theta|\mathbf{y}) = \frac{p(\mathbf{y}|\theta)\,p(\theta)}{p(\mathbf{y})}$$

1950

- Modèles simples
- Lois a priori non informatives
- Lois a priori conjuguées
- Solutions explicites

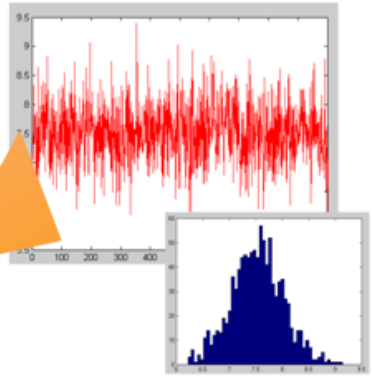
$$p(\theta_1|y) = \iiint p(\theta_1, \theta_2, \ldots, \theta_n|y)\, d\theta_2 \ldots d\theta_n$$

- Analyse numérique
- Modèles simples

1990

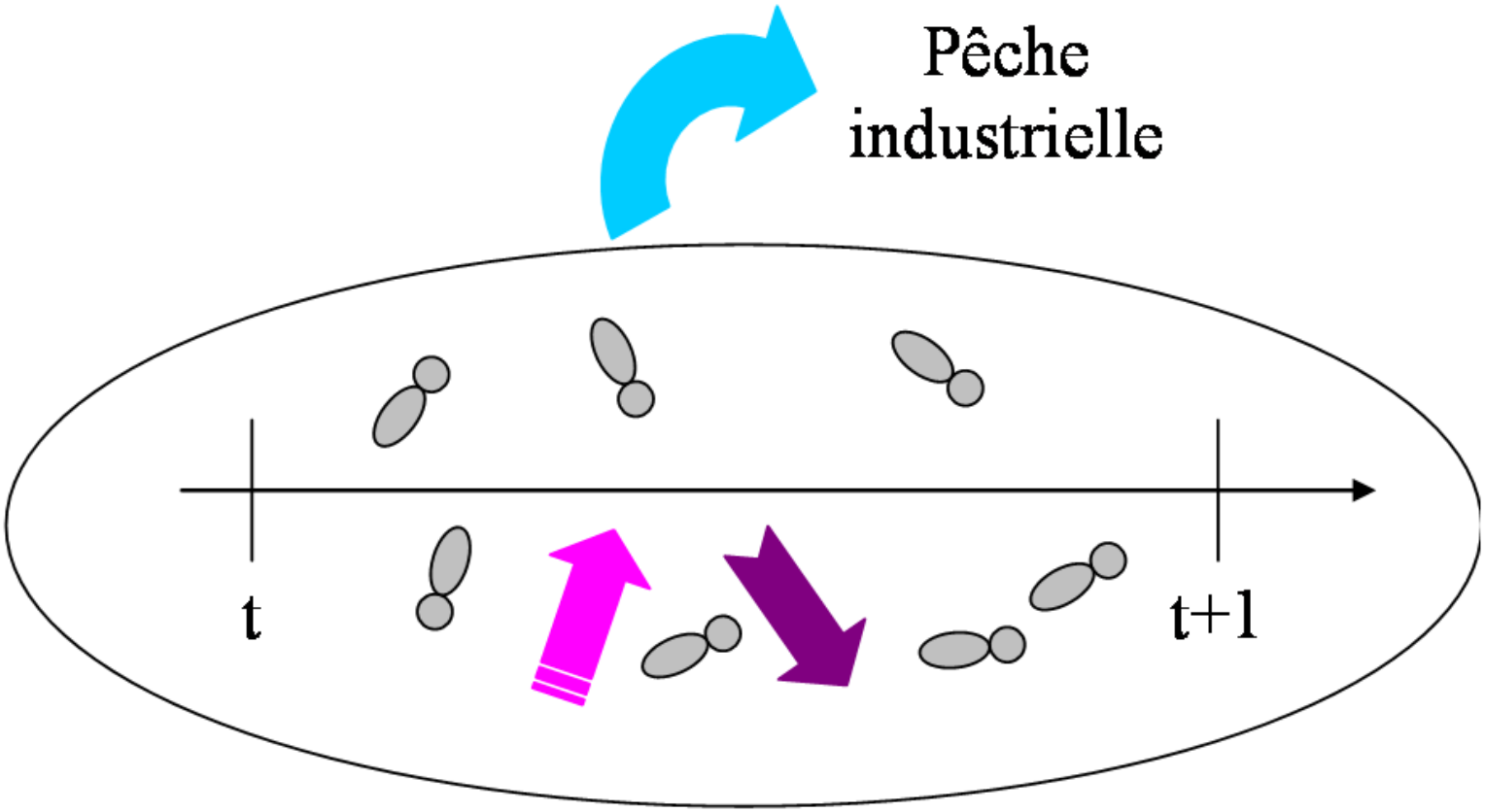
MCMC

- Gibbs sampling
- Hasting-Metropolis
- Autres



Effort de pêche optimal

Permettre le renouvellement naturel d'un stock de poissons



Effort de pêche optimal

Le Germon (thon blanc) de l'Atlantique Sud

- Polacheck, T., Hilborn, R., and Punt, A.E. 1993. *Fitting surplus production models: comparing methods and measuring uncertainty.* Can. J. Fish. Aquat. Sci. 50: 2597–2607.

	0	1	2	3	4	5	6	7	8	9
c_t								15.9	25.7	28.5
y_t								61.9	79.0	55.6
c_t	23.7	25	33.3	28.2	19.7	17.5	19.3	21.6	23.1	22.5
y_t	44.6	56.9	38.3	33.8	36.1	42.0	36.6	36.3	38.8	34.3
c_t	22.5	23.6	29.1	14.4	13.2	28.4	34.6	37.5	25.9	25.3
y_t	37.6	34.0	32.2	26.9	36.6	30.1	30.8	23.4	22.3	21.9

Effort de pêche optimal

Le Germon (thon blanc) de l'Atlantique Sud

- Polacheck, T., Hilborn, R., and Punt, A.E. 1993. *Fitting surplus production models: comparing methods and measuring uncertainty.* Can. J. Fish. Aquat. Sci. 50: 2597–2607.
  - Les captures industrielles c_t sont en milliers de tonnes

	0	1	2	3	4	5	6	7	8	9
c_t								15.9	25.7	28.5
y_t								61.9	79.0	55.6
c_t	23.7	25	33.3	28.2	19.7	17.5	19.3	21.6	23.1	22.5
y_t	44.6	56.9	38.3	33.8	36.1	42.0	36.6	36.3	38.8	34.3
c_t	22.5	23.6	29.1	14.4	13.2	28.4	34.6	37.5	25.9	25.3
y_t	37.6	34.0	32.2	26.9	36.6	30.1	30.8	23.4	22.3	21.9

Effort de pêche optimal

Le Germon (thon blanc) de l'Atlantique Sud

- Polacheck, T., Hilborn, R., and Punt, A.E. 1993. *Fitting surplus production models: comparing methods and measuring uncertainty.* Can. J. Fish. Aquat. Sci. 50: 2597–2607.
  - Les captures industrielles c_t sont en milliers de tonnes
  - Les observations y_t sont en kg de biomasse par 100 hameçons

	0	1	2	3	4	5	6	7	8	9
c_t								15.9	25.7	28.5
y_t								61.9	79.0	55.6
c_t	23.7	25	33.3	28.2	19.7	17.5	19.3	21.6	23.1	22.5
y_t	44.6	56.9	38.3	33.8	36.1	42.0	36.6	36.3	38.8	34.3
c_t	22.5	23.6	29.1	14.4	13.2	28.4	34.6	37.5	25.9	25.3
y_t	37.6	34.0	32.2	26.9	36.6	30.1	30.8	23.4	22.3	21.9

Schaefer surplus production model

Définition de l'effort optimal de pêche

- Fondé sur deux paramètres du modèle décrit ci-après :

$$E_{MSP}|\alpha, \beta = \frac{\alpha}{2\beta}$$

Schaefer surplus production model

Définition de l'effort optimal de pêche

- Fondé sur deux paramètres du modèle décrit ci-après :

$$E_{MSP}|\alpha, \beta = \frac{\alpha}{2\beta}$$

- Modèle de processus

$$\begin{aligned} B_{t+1}|\mu_t, \tau_B &\sim dlnorm\left(\ln \mu_t, \tau_B\right) \\ \mu_t &\equiv \max\left(B_t + \alpha B_t\left(1 - \frac{B_t}{\kappa}\right) - c_t, \varepsilon\right) \end{aligned}$$

Schaefer surplus production model

Définition de l'effort optimal de pêche

- Fondé sur deux paramètres du modèle décrit ci-après :

$$E_{MSP}|\alpha, \beta = \frac{\alpha}{2\beta}$$

- Modèle de processus

$$\begin{aligned} B_{t+1}|\mu_t, \tau_B &\sim dlnorm\left(\ln \mu_t, \tau_B\right) \\ \mu_t &\equiv \max\left(B_t + \alpha B_t\left(1 - \frac{B_t}{\kappa}\right) - c_t, \varepsilon\right) \end{aligned}$$

- Modèle d'observable

$$Y_t|\beta, B_t, \tau_Y \sim dlnorm\left(\ln \beta B_t, \tau_Y\right)$$

Schaefer surplus production model

Définition de l'effort optimal de pêche

- Fondé sur deux paramètres du modèle décrit ci-après :

$$E_{MSP}|\alpha, \beta = \frac{\alpha}{2\beta}$$

- Modèle de processus

$$\begin{aligned} B_{t+1}|\mu_t, \tau_B &\sim dlnorm\left(\ln \mu_t, \tau_B\right) \\ \mu_t &\equiv \max\left(B_t + \alpha B_t\left(1 - \frac{B_t}{\kappa}\right) - c_t, \varepsilon\right) \end{aligned}$$

- Modèle d'observable

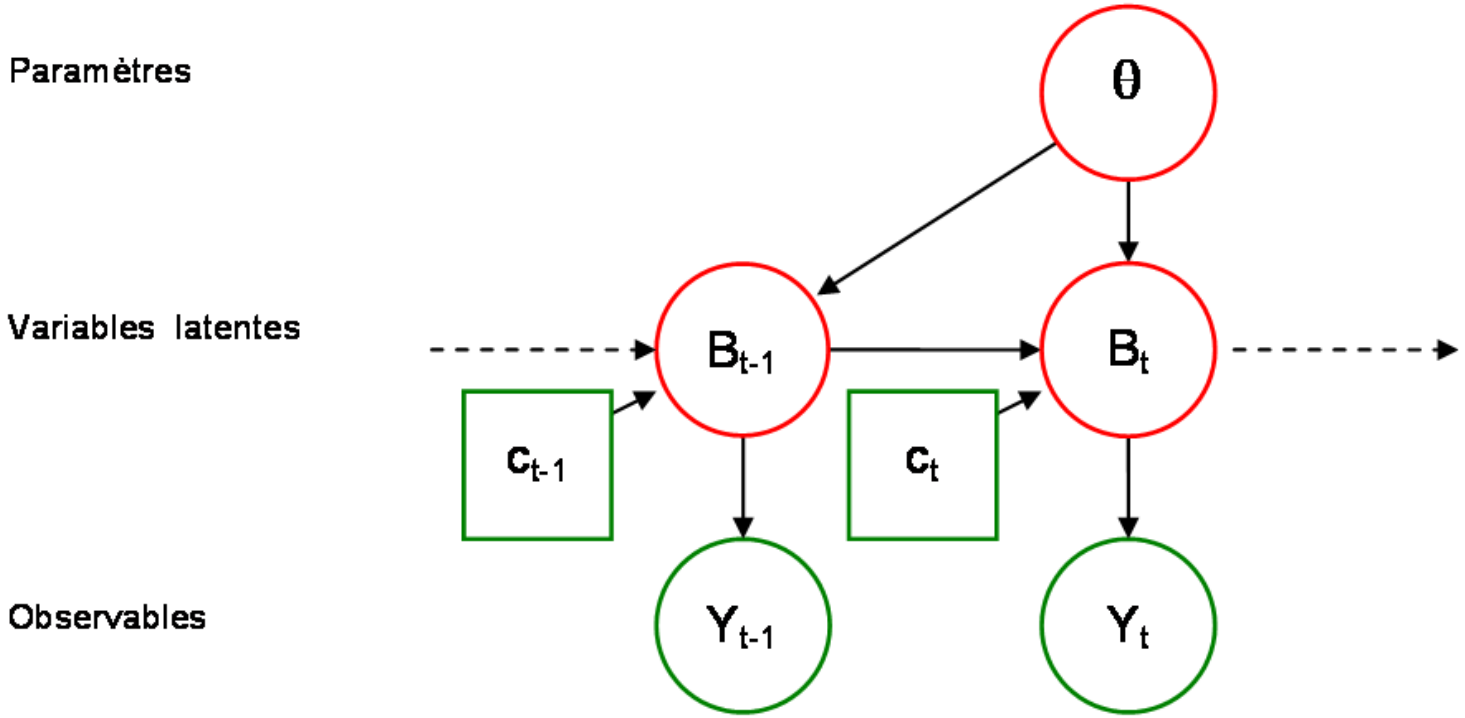
$$Y_t|\beta, B_t, \tau_Y \sim dlnorm\left(\ln \beta B_t, \tau_Y\right)$$

- Schaefer, M.B. 1954. *Some aspects of the dynamics of populations important to the management of commercial marine fisheries.* Inter-Am. Trop. Tuna Comm. Bull. 1: 27–56.

Modélisation hiérarchique

Modèle d'espace d'états non linéaire

Meyer R. and Millar R.B. 1999. *BUGS in Bayesian stock assessments.* Can. J. Fish. Aquat. Sci. 56: 1078–1086 (1999)



Les éléments du problème

Quelsques hypothèses

- $B_t \geq 0$ représente la biomasse disponible au 01 janvier de l'année t

Les éléments du problème

Quelsques hypothèses

- $B_t \geq 0$ représente la biomasse disponible au 01 janvier de l'année t
  - C'est une quantitée inobservable (variable latente).

Les éléments du problème

Quelsques hypothèses

- $B_t \geq 0$ représente la biomasse disponible au 01 janvier de l'année t
  - C'est une quantitée inobservable (variable latente).
  - dB_t / dt est le taux naturel de renouvellement durant l'année t

Les éléments du problème

Quelsques hypothèses

- $B_t \geq 0$ représente la biomasse disponible au 01 janvier de l'année t
  - C'est une quantitée inobservable (variable latente).
  - dB_t / dt est le taux naturel de renouvellement durant l'année t
    - Par hypothèse

$$\frac{dB_t}{dt} \propto B_t$$

Les éléments du problème

Quelsques hypothèses

- $B_t \geq 0$ représente la biomasse disponible au 01 janvier de l'année t
  - C'est une quantitée inobservable (variable latente).
  - dB_t / dt est le taux naturel de renouvellement durant l'année t
    - Par hypothèse

$$\frac{dB_t}{dt} \propto B_t$$

- c_t représente les captures industrielles durant l'année calendaire t

Les éléments du problème

Quelsques hypothèses

- $B_t \geq 0$ représente la biomasse disponible au 01 janvier de l'année t
  - C'est une quantitée inobservable (variable latente).
  - dB_t / dt est le taux naturel de renouvellement durant l'année t
    - Par hypothèse

$$\frac{dB_t}{dt} \propto B_t$$

- c_t représente les captures industrielles durant l'année calendaire t
  - Par hypothèse, c_t est une quantité certaine

Les éléments du problème

Quelsques hypothèses

- $B_t \geq 0$ représente la biomasse disponible au 01 janvier de l'année t
  - C'est une quantitée inobservable (variable latente).
  - dB_t / dt est le taux naturel de renouvellement durant l'année t
    - Par hypothèse

$$\frac{dB_t}{dt} \propto B_t$$

- c_t représente les captures industrielles durant l'année calendaire t
  - Par hypothèse, c_t est une quantité certaine
- L'observable, Y_t, est proportionnelle à la biomasse cachée

$$Y_t = qB_t$$

Évolution naturelle de la biomasse associée à une population de poissons

- La variation de la biomasse sur l'intervalle dt est proportionnelle à la biomasse initiale

$$\frac{dB_t}{dt} = f\left(B_t\right) \times B_t$$

Évolution naturelle de la biomasse associée à une population de poissons

- La variation de la biomasse sur l'intervalle dt est proportionnelle à la biomasse initiale

$$\frac{dB_t}{dt} = f\left(B_t\right) \times B_t$$

- Soit $n\left(B_t\right)$ et $m\left(B_t\right)$ les taux de natalité et de mortalité sur l'intervalle dt

$$f\left(B_t\right) = n\left(B_t\right) - m\left(B_t\right)$$

Évolution naturelle de la biomasse associée à une population de poissons

- La variation de la biomasse sur l'intervalle dt est proportionnelle à la biomasse initiale

$$\frac{dB_t}{dt} = f\left(B_t\right) \times B_t$$

- Soit $n\left(B_t\right)$ et $m\left(B_t\right)$ les taux de natalité et de mortalité sur l'intervalle dt

$$f\left(B_t\right) = n\left(B_t\right) - m\left(B_t\right)$$

- Existence de facteurs limitants

$$B_t \nearrow \quad \Rightarrow \left\{ \begin{array}{l} n\left(B_t\right) \searrow \\ m\left(B_t\right) \nearrow \end{array} \right. \Leftrightarrow f\left(B_t\right) \searrow$$

Évolution naturelle de la biomasse associée à une population de poissons

- La variation de la biomasse sur l'intervalle dt est proportionnelle à la biomasse initiale
$$\frac{dB_t}{dt} = f\left(B_t\right) \times B_t$$
- Soit $n\left(B_t\right)$ et $m\left(B_t\right)$ les taux de natalité et de mortalité sur l'intervalle dt
$$f\left(B_t\right) = n\left(B_t\right) - m\left(B_t\right)$$
- Existence de facteurs limitants
$$B_t \nearrow \quad \Rightarrow \left\{ \begin{array}{l} n\left(B_t\right) \searrow \\ m\left(B_t\right) \nearrow \end{array} \right. \Leftrightarrow f\left(B_t\right) \searrow$$
- Une fonction simple
$$f\left(B_t\right) = \alpha\left(1 - \frac{B_t}{\kappa}\right), \quad \alpha, \kappa > 0$$

Équation de Verhulst

- L'évolution naturelle de la biomasse est une équation différentielle ordinaire non linéaire connue sous le nom *d'équation de Verhulst*

$$\frac{dB_t}{dt} = \alpha \left(1 - \frac{B_t}{\kappa}\right) \times B_t$$

Équation de Verhulst

- L'évolution naturelle de la biomasse est une équation différentielle ordinaire non linéaire connue sous le nom *d'équation de Verhulst*

$$\frac{dB_t}{dt} = \alpha \left(1 - \frac{B_t}{\kappa}\right) \times B_t$$

- Sa solution est une courbe de croissance logistique

$$B_t = \frac{\kappa}{1 + c\kappa \exp\left(-\alpha t\right)}$$

Équation de Verhulst

- L'évolution naturelle de la biomasse est une équation différentielle ordinaire non linéaire connue sous le nom *d'équation de Verhulst*

$$\frac{dB_t}{dt} = \alpha \left(1 - \frac{B_t}{\kappa}\right) \times B_t$$

- Sa solution est une courbe de croissance logistique

$$B_t = \frac{\kappa}{1 + c\kappa \exp\left(-\alpha t\right)}$$

  - On a

$$B_0 = \frac{\kappa}{1+c\kappa}, \quad B_\infty = \kappa$$

Évolution naturelle de la biomasse (suite)

Maximum surplus production

- On a

$$\frac{dB_t}{dt} = f \times B_t = \alpha B_t \left(1 - \frac{B_t}{\kappa}\right) \equiv g\left(B_t\right)$$

Évolution naturelle de la biomasse (suite)

Maximum surplus production

- On a

$$\frac{dB_t}{dt} = f \times B_t = \alpha B_t \left(1 - \frac{B_t}{\kappa}\right) \equiv g\left(B_t\right)$$

- C'est un trinôme du second degré

Évolution naturelle de la biomasse (suite)

Maximum surplus production

- On a

$$\frac{dB_t}{dt} = f \times B_t = \alpha B_t \left(1 - \frac{B_t}{\kappa}\right) \equiv g\left(B_t\right)$$

- C'est un trinôme du second degré
  - $g\left(B_t\right) = 0 \Longleftrightarrow B_t \in \{0, \kappa\}$

Évolution naturelle de la biomasse (suite)

Maximum surplus production

- On a

$$\frac{dB_t}{dt} = f \times B_t = \alpha B_t \left(1 - \frac{B_t}{\kappa}\right) \equiv g\left(B_t\right)$$

- C'est un trinôme du second degré
  - $g\left(B_t\right) = 0 \Longleftrightarrow B_t \in \{0, \kappa\}$
  - $B_t = \kappa/2 \equiv B_{MSP} \Longleftrightarrow g\left(B_{MSP}\right) = \frac{\alpha\kappa}{4} \equiv MSP$

Évolution naturelle de la biomasse (suite)

Maximum surplus production

- On a
$$\frac{dB_t}{dt} = f \times B_t = \alpha B_t \left(1 - \frac{B_t}{\kappa}\right) \equiv g\left(B_t\right)$$
- C'est un trinôme du second degré
  - $g\left(B_t\right) = 0 \Longleftrightarrow B_t \in \{0, \kappa\}$
  - $B_t = \kappa/2 \equiv B_{MSP} \Longleftrightarrow g\left(B_{MSP}\right) = \frac{\alpha\kappa}{4} \equiv MSP$
- Le facteur $1 - \frac{B_t}{\kappa}$ apparaît comme un frein à la croissance exponentielle

Évolution naturelle de la biomasse (suite)

Maximum surplus production

- On a
$$\frac{dB_t}{dt} = f \times B_t = \alpha B_t \left(1 - \frac{B_t}{\kappa}\right) \equiv g\left(B_t\right)$$
- C'est un trinôme du second degré
  - $g\left(B_t\right) = 0 \Longleftrightarrow B_t \in \{0, \kappa\}$
  - $B_t = \kappa/2 \equiv B_{MSP} \Longleftrightarrow g\left(B_{MSP}\right) = \frac{\alpha\kappa}{4} \equiv MSP$
- Le facteur $1 - \frac{B_t}{\kappa}$ apparaît comme un frein à la croissance exponentielle
  - En effet, sans ce facteur :
$$\frac{dB_t}{dt} = \alpha B_t \Longrightarrow B_t = B_0 \exp\left(\alpha t\right)$$

Le modèle

Partie déterministe

- Modéliser le processus

$$B_{t+1} = B_t + \alpha B_t \left(1 - \frac{B_t}{\kappa}\right) - c_t$$

Le modèle

Partie déterministe

- Modéliser le processus

$$B_{t+1} = B_t + \alpha B_t \left(1 - \frac{B_t}{\kappa}\right) - c_t$$

  - La biomasse B_t est une variable aléatoire cachée.

Le modèle

Partie déterministe

- Modéliser le processus

$$B_{t+1} = B_t + \alpha B_t \left(1 - \frac{B_t}{\kappa}\right) - c_t$$

  - La biomasse B_t est une variable aléatoire cachée.

- Modéliser l'observable

Le modèle

Partie déterministe

- Modéliser le processus

$$B_{t+1} = B_t + \alpha B_t \left(1 - \frac{B_t}{\kappa}\right) - c_t$$

  - La biomasse B_t est une variable aléatoire cachée.

- Modéliser l'observable

  - L'observable, Y_t, est le poids des poissons rapporté au nombre d'hameçons sur la ligne de pêche.

Le modèle

Partie déterministe

- Modéliser le processus

$$B_{t+1} = B_t + \alpha B_t \left(1 - \frac{B_t}{\kappa}\right) - c_t$$

  - La biomasse B_t est une variable aléatoire cachée.

- Modéliser l'observable
  - L'observable, Y_t, est le poids des poissons rapporté au nombre d'hameçons sur la ligne de pêche.
  - L'observable est proportionnelle à la biomasse :

$$Y_t = \beta B_t$$

L'effort de pêche optimal

But : préserver le stock

- L'observable correspondant au taux naturel de renouvellement maximum

$$\left.\begin{array}{r} Y_t = \beta B_t \\ B_{MSP} = \frac{\kappa}{2} \end{array}\right\} \Rightarrow Y_{MSP} = \frac{\beta\kappa}{2}$$

L'effort de pêche optimal

But : préserver le stock

- L'observable correspondant au taux naturel de renouvellement maximum

$$\left.\begin{array}{r} Y_t = \beta B_t \\ B_{MSP} = \frac{\kappa}{2} \end{array}\right\} \Rightarrow Y_{MSP} = \frac{\beta\kappa}{2}$$

- *Schaefer (1954)* définit l'effort de pêche optimal comme le quotient entre le MSP et cet indice

$$\left.\begin{array}{r} MSP = \frac{\alpha\kappa}{4} \\ Y_{MSP} = \frac{\beta\kappa}{2} \end{array}\right\} \Rightarrow E_{MSP} = \frac{MSP}{Y_{MSP}} = \frac{\alpha}{2\beta}$$

Le hasard intervient

Bruiter les processus

- Bruiter le modèle de processus (bruit blanc)

$$B_t = \left(B_{t-1} + \alpha B_{t-1} \left(1 - \frac{B_{t-1}}{\kappa} \right) - C_t \right) \exp\left(u_t \right)$$

Le hasard intervient

Bruiter les processus

- Bruiter le modèle de processus (bruit blanc)

$$B_t = \left(B_{t-1} + \alpha B_{t-1} \left(1 - \frac{B_{t-1}}{\kappa} \right) - C_t \right) \exp\left(u_t \right)$$

- Bruiter le modèle d'observable (bruit blanc)

$$Y_t = q B_t \exp\left(\varepsilon_t \right)$$

Le hasard intervient

Bruiter les processus

- Bruiter le modèle de processus (bruit blanc)

$$B_t = \left(B_{t-1} + \alpha B_{t-1} \left(1 - \frac{B_{t-1}}{\kappa} \right) - C_t \right) \exp\left(u_t \right)$$

- Bruiter le modèle d'observable (bruit blanc)

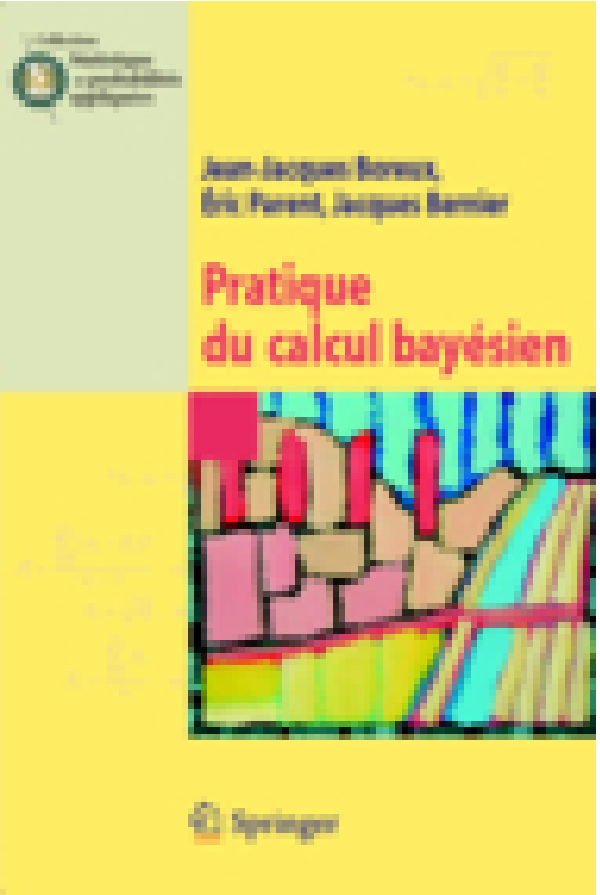
$$Y_t = q B_t \exp\left(\varepsilon_t \right)$$

- Bruiter les deux modèles

Résultats

WinBUGS

	Q_5	Q_{50}	Q_{95}
κ	176	269	372
α	0.17	0.29	0.51
β	0.0027	0.0037	0.0057
E_{MSP}	15	40	95
B_{1990}	46	80	123

Un peu de pub

Notre dernier livre