

Distributions *a priori*

Olivier David

MaIAGE, INRA, Jouy-en-Josas

BioBayes, Pornic, mars 2019

1. Exemple

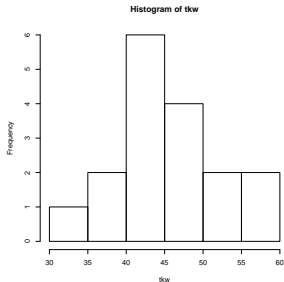
Essai à la ferme réalisé en 2015 pour comparer 14 populations de blé, comprenant 17 parcelles.

Essai réalisé en collaboration avec l'équipe "Diversité, Evolution et Adaptation des Populations" de l'UMR "Génétique Quantitative et Évolution-Le Moulon" dans le cadre d'un projet sur le développement de l'agriculture biologique pour le blé.

Plan de l'essai :

10	13	8	3	1	12	14	5	4
9	11	4	12	2	7	13	6	

Variable poids de mille grains (tkw) en g.



L'observation y_{ij} pour la population i et la répétition j est modélisée par :

$$y_{ij} = \alpha_i + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim \mathcal{N}(0, \sigma_\varepsilon^2).$$

Il y a 15 paramètres et 17 données.

Anova

Analysis of Variance Table

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
germplasm	14	36691	2620.82	223.73	0.0004325 ***
Residuals	3	35	11.71		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

	Estimate	Std. Error	t value	Pr(> t)
germplasm 10	58.410	3.423	17.066	0.000438 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.423 on 3 degrees of freedom
Multiple R-squared: 0.999, Adjusted R-squared: 0.9946
F-statistic: 223.7 on 14 and 3 DF, p-value: 0.0004325

	2.5 %	97.5 %
germplasm 10	47.51768	69.30232

Quelles lois *a priori* utiliser pour α_i et σ_ϵ ?

2. Loi *a priori* informative

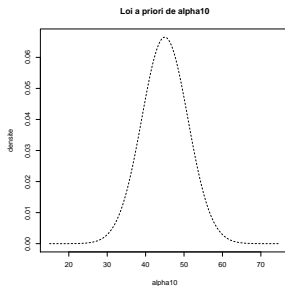
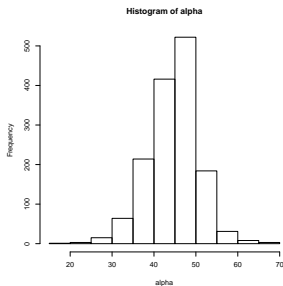
Loi *a priori informative* : loi qui cherche à prendre en compte l'information *a priori* dont on dispose sur les paramètres.

On utilise les résultats de l'anova : $\alpha_{10} \sim \mathcal{N}(58.410, 3.423^2)$.

	Mean	SD
alpha[10]	58.398365	2.0805987
sigma	2.649136	0.5259157

On peut construire une loi *a priori* informative en utilisant des données venant d'autres expériences, de la littérature etc.

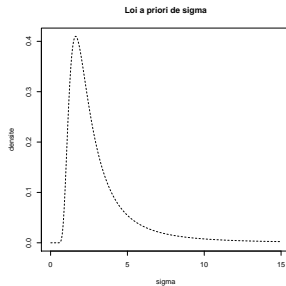
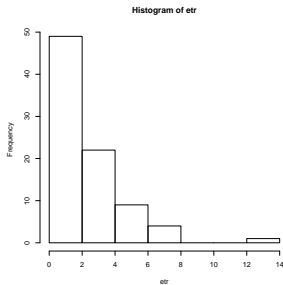
Histogramme des $\hat{\alpha}_i$ dans d'autres essais et loi *a priori* de α_{10} :



Loi *a priori* normale pour α : $\alpha_i \sim \mathcal{N}(\mu_0, \sigma_{\alpha 0}^2)$, $i = 1, \dots, 14$.

Moyenne des $\hat{\alpha}_i \approx 45 \Rightarrow \mu_0 = 45$. Ecart-type des $\hat{\alpha}_i \approx 6 \Rightarrow \sigma_{\alpha 0} = 6$.

Histogramme des $\hat{\sigma}_\epsilon$ dans d'autres essais et loi *a priori* de σ_ϵ :

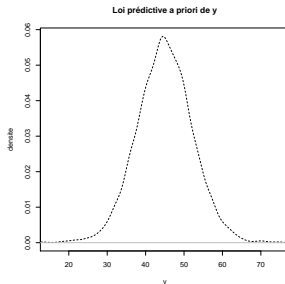
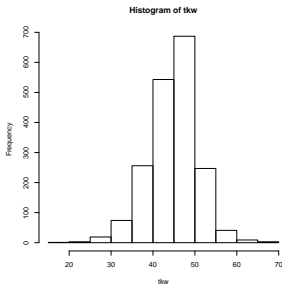


Loi *a priori* gamma pour σ_ϵ^{-2} : $\sigma_\epsilon^{-2} \sim \mathcal{G}(1, 4)$.

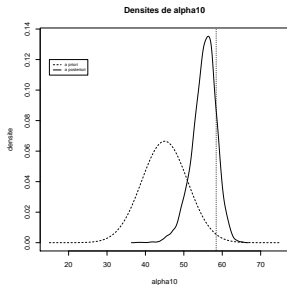
Loi prédictive *a priori* de y_{101} :

$$p(y_{101}) = \int p(y_{101} | \alpha_{10}, \sigma_\varepsilon) p(\alpha_{10}) p(\sigma_\varepsilon) d\alpha_{10} d\sigma_\varepsilon.$$

Histogramme de tkw dans d'autres essais et loi prédictive *a priori* de y_{101} :



Lois de α_{10} :



Résumés de la loi *a posteriori* :

	2.5%	25%	50%	75%	97.5%
alpha[10]	48.152562	53.36665	55.55667	57.411518	60.893645
sigma	1.860142	2.55137	3.10284	3.841427	5.898445

L'intervalle de crédibilité de α_{10} est moins grand que celui de l'anova.

3. Loi *a priori* “non informative”

Loi *a priori* “non informative” : loi qui cherche à apporter le moins d'information *a priori* possible sur les paramètres.

Loi *a priori* normale pour $\alpha|\sigma_\varepsilon$: $\alpha_i|\sigma_\varepsilon \sim \mathcal{N}\left(\mu_0, \frac{\sigma_\varepsilon^2}{r_0}\right)$, $i = 1, \dots, 14$.

Loi *a posteriori* de $\alpha|\sigma_\varepsilon$:

$$\alpha_i|\sigma_\varepsilon, y \sim \mathcal{N}\left(\mu'_i, \frac{\sigma_\varepsilon^2}{r'_i}\right), \quad i = 1, \dots, 14,$$
$$\mu'_i = \frac{r_0\mu_0 + r_i y_{i\bullet}}{r_0 + r_i}, \quad r'_i = r_0 + r_i,$$

r_i est le nombre de répétitions de la population i ,

$y_{i\bullet}$ est la moyenne des observations pour la population i .

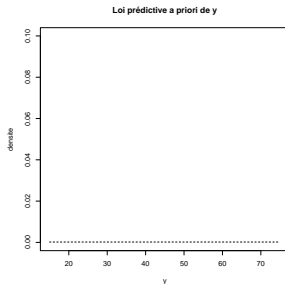
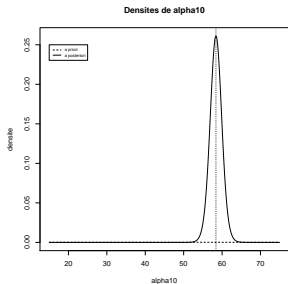
Loi *a priori conjuguée* avec la vraisemblance : la loi *a posteriori* appartient à la même famille de lois que la loi *a priori*.

La loi *a priori* apporte la même information qu'un jeu de données comprenant r_0 répétitions de la population i et tel que $\hat{\alpha}_i = \mu_0$.

La loi *a priori* a peu d'influence sur les résultats si les répétitions sont nombreuses.

Loi "non informative" pour α : $r_0 = 10^{-6}$.

Lois de α_{10} et loi prédictive *a priori* de y_{101} :



La loi *a priori* de α_{10} a une grande variance, elle est vague.

Résumés de la loi *a posteriori* :

	2.5%	25%	50%	75%	97.5%
alpha[10]	55.235266	57.380575	58.400574	59.413355	61.524943
sigma	1.144239	1.377826	1.536445	1.726912	2.240271

Il est en général plus difficile de construire une loi *a priori* “non informative” qu’une loi informative.

Dans cette école, on se limite à des lois vagues.

Une analyse de sensibilité à la loi *a priori* peut aider à détecter des problèmes.

Quand $r_0 \rightarrow 0$, la loi *a priori* de $\alpha|\sigma$ tend vers la loi uniforme sur \mathbb{R} .

Loi *a priori impropre* : loi *a priori* dont l'intégrale ne vaut pas 1 mais $+\infty$; elle n'est pas une véritable loi de probabilité.

Dans ce cas, il faut vérifier si la loi *a posteriori* est propre.

Quand $r_0 \rightarrow 0$, la loi *a posteriori* de $\alpha|\sigma$ est propre :

$$\alpha_i|\sigma_\varepsilon, y \sim \mathcal{N}\left(\alpha'_i, \frac{\sigma_\varepsilon^2}{r'_i}\right), \quad i = 1, \dots, 14,$$
$$\alpha'_i = y_{i\bullet}, \quad r'_i = r_i.$$

On peut donc utiliser la loi *a priori* uniforme pour $\alpha|\sigma_\varepsilon$.

4. Loi *a priori* hiérarchique

Loi *a priori* hiérarchique : loi *a priori* qui a plusieurs niveaux.

Loi *a priori* de α à 2 niveaux :

$$\text{Niveau 1} : y_{ij} \sim \mathcal{N}(\alpha_i, \sigma_\varepsilon^2),$$

$$\text{Niveau 2} : \alpha_i \sim \mathcal{N}(\mu, \sigma_\alpha^2), \quad \sigma_\varepsilon^{-2} \sim \mathcal{G}(1, 4),$$

$$\text{Niveau 3} : \mu \sim \mathcal{N}(0, \sigma_\mu^2), \quad \sigma_\alpha^{-2} \sim \mathcal{G}(\gamma, \gamma), \\ \sigma_\mu \rightarrow +\infty, \quad \gamma \rightarrow 0.$$

μ et σ_α sont des paramètres inconnus du niveau 2, appelés *hyperparamètres*.

On ne peut pas utiliser cette loi car la loi *a posteriori* est impropre.

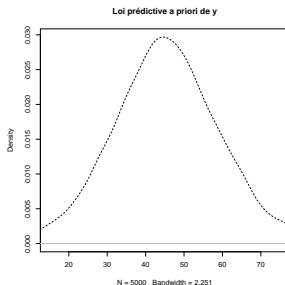
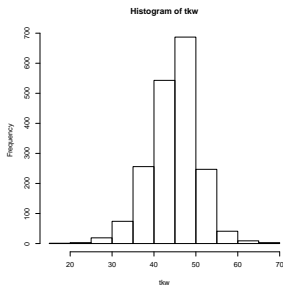
Loi *a priori* de α à 2 niveaux :

$$\text{Niveau 1} : y_{ij} \sim \mathcal{N}(\alpha_i, \sigma_\varepsilon^2),$$

$$\text{Niveau 2} : \alpha_i \sim \mathcal{N}(\mu, \sigma_\alpha^2), \quad \sigma_\varepsilon^{-2} \sim \mathcal{G}(1, 4),$$

$$\text{Niveau 3} : \mu \sim \mathcal{N}(45, 10^2), \quad \sigma_\alpha \sim \mathcal{N}^+(0, 10^2).$$

Histogramme de tkw dans d'autres essais et loi prédictive *a priori* de y_{101} :



Résumés de la loi *a posteriori* :

	2.5%	25%	50%	75%	97.5%
alpha[10]	45.097221	52.442597	55.387277	57.598797	61.534847
mu	42.097967	44.703527	45.897282	47.163301	49.796280
sigma	1.839705	2.610050	3.278263	4.293542	7.188869
sigma.alpha	1.723310	5.089808	6.239840	7.462387	10.520161

L'estimation de α_{10} s'est rapprochée de μ : on dit qu'elle est *rétrécie*.

L'intervalle de crédibilité de α_{10} est moins grand que celui de l'anova.

5. Références bibliographiques

- ▶ George E. P. Box and George C. Tiao. Bayesian inference in statistical analysis. Addison-Wesley, 1973.
- ▶ Collectif BioBayes. Initiation à la Statistique Bayésienne. Chapitre 7. Ellipses, 2015.
- ▶ Jean-Jacques Dreesbeke, Jeanne Fine, and Gilbert Saporta, editors. Méthodes bayésiennes en statistique. Editions Technip, 2002.
- ▶ A. Gelman (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). Bayesian analysis, 1(3), 515-534.
- ▶ Jean-Michel Marin and Christian Robert. Bayesian core : a practical approach to computational Bayesian statistics. Springer Science & Business Media, 2007.
- ▶ Christian P. Robert. Le choix bayésien : Principes et pratique. Springer, 2006.