INSEA

61

MCMC diagnostics in 1D

Metrics: R hat de Gelman-Rubin

Moyenne et variance de chaque chaîne

$$\text{Chaîne 1} : X_1^{(1)}, \ldots, X_n^{(1)} \longrightarrow \overline{\mathbf{X}}^{(1)} \quad , \quad \sigma^{\mathbf{2}(1)}$$

$$\vdots$$

$$\text{Chaîne j} : X_1^{(j)}, \ldots, X_n^{(j)} \longrightarrow \overline{\mathbf{X}}^{(j)} \quad , \quad \sigma^{\mathbf{2}(j)}$$

$$\vdots$$

$$\text{Chaîne m} : X_1^{(m)}, \ldots, X_n^{(m)} \longrightarrow \overline{\mathbf{X}}^{(m)} \quad , \quad \sigma^{\mathbf{2}(m)}$$

Moyenne des moyennes $\overline{\overline{\mathbf{X}}}$

Moyenne des variances “within” $$W = \frac{1}{m}\sum_{j=1}^{m} \sigma^{\mathbf{2}(j)}$$

Variance des moyennes “between” $$B = \frac{n}{m-1}\sum_{j}^{m}(\overline{\mathbf{X}}^{(j)} - \overline{\overline{\mathbf{X}}})^2$$

Si les chaînes convergent vers la même moyenne:

$$\frac{B}{n} \to 0$$

“R hat” de Gelman-Rubin (1992)

$$\widehat{R} \stackrel{\text{def}}{=} \sqrt{\frac{\frac{n-1}{n}W + \frac{1}{n}B}{W}} \to r \begin{cases} = 1 \text{ si } \dfrac{B}{n} \to 0 \\ > 1 \text{ sinon} \end{cases}$$

Sinon $\dfrac{B}{n} \to b^\star > 0$

Détails en TD

$$\widehat{R} \stackrel{\text{def}}{=} \sqrt{\frac{\frac{n-1}{n}\sigma^2_{within} + \frac{1}{n}\sigma^2_{between}}{\sigma^2_{within}}}$$

En pratique on veut R < 1.01

Pourquoi n ?

car $\mathbb{V}(\overline{\mathbf{X}}^{(j)}) = \frac{\sigma^2}{n}$

Metrics: R hat de Gelman-Rubin

Moyenne et variance de chaque chaîne

$\text{Chaîne 1} : X_1^{(1)}, \dots, X_n^{(1)} \longrightarrow \overline{\mathbf{X}}^{(1)} \quad , \quad \sigma^{\mathbf{2}(1)}$

$\vdots$

$\text{Chaîne j} : X_1^{(j)}, \dots, X_n^{(j)} \longrightarrow \overline{\mathbf{X}}^{(j)} \quad , \quad \sigma^{\mathbf{2}(j)}$

$\vdots$

$\text{Chaîne m} : X_1^{(m)}, \dots, X_n^{(m)} \longrightarrow \overline{\mathbf{X}}^{(m)} \quad , \quad \sigma^{\mathbf{2}(m)}$

Moyenne des moyennes $\overline{\overline{\mathbf{X}}}$

Variance des moyennes "between" $B = \frac{n}{m-1} \sum_{j}^{m} (\overline{\mathbf{X}}^{(j)} - \overline{\overline{\mathbf{X}}})^2$

Moyenne des variances "within" $W = \frac{1}{m} \sum_{j=1}^{m} \sigma^{\mathbf{2}(j)}$

Pourquoi n ?

car $\mathbb{V}(\overline{\mathbf{X}}^{(j)}) = \frac{\sigma^2}{n}$

Si les chaînes convergent vers la même moyenne: $\frac{B}{n} \to 0$

Sinon $\frac{B}{n} \to b^\star > 0$

"R hat" de Gelman-Rubin (1992)

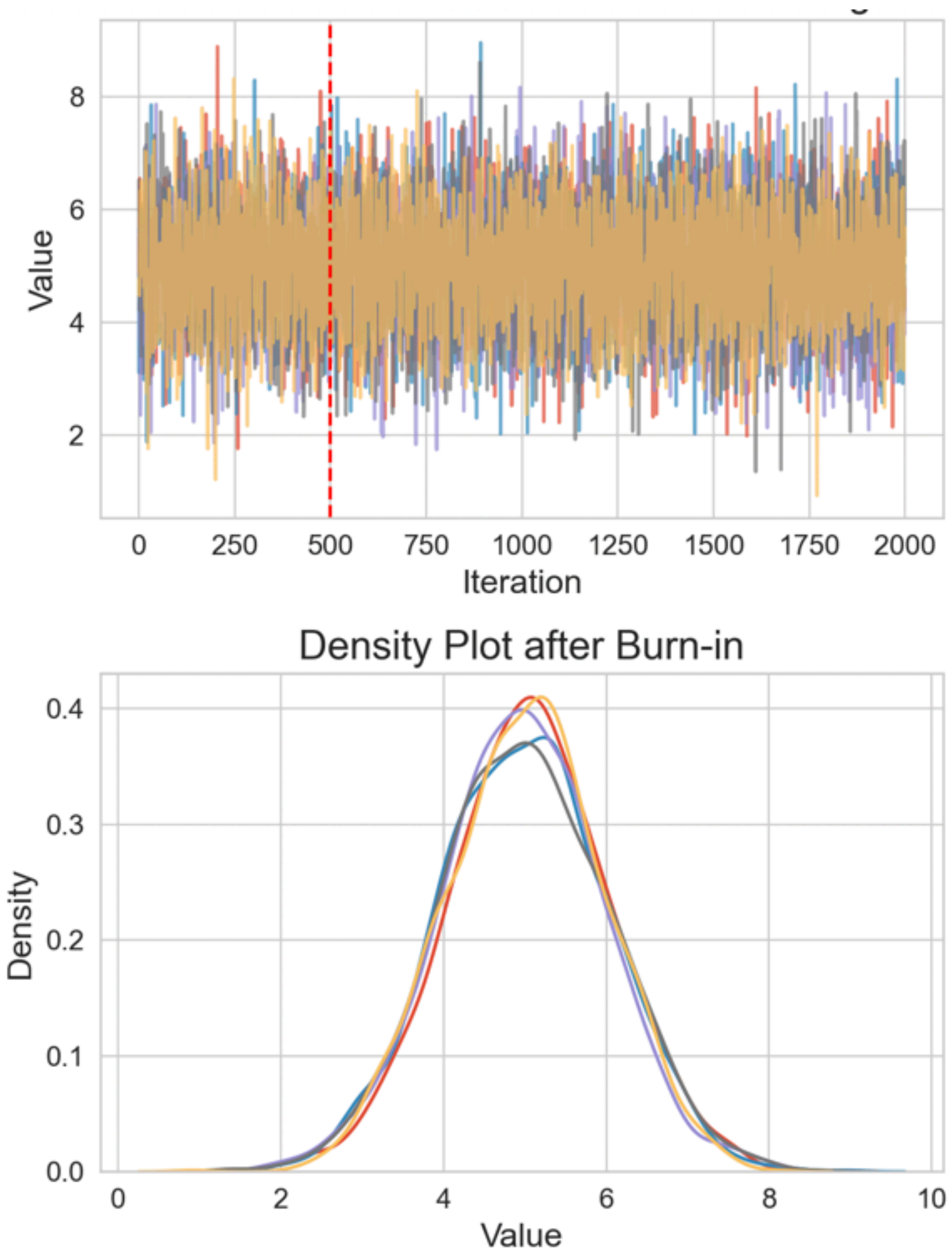
$$\widehat{R} \stackrel{\text{def}}{=} \sqrt{\frac{\frac{n-1}{n} W + \frac{1}{n} B}{W}} \to r \begin{cases} = 1 \text{ si } \frac{B}{n} \to 0 \\ > 1 \text{ sinon} \end{cases}$$

$$\widehat{R} \stackrel{\text{def}}{=} \sqrt{\frac{\frac{n-1}{n}\sigma^2_{within} + \frac{1}{n}\sigma^2_{between}}{\sigma^2_{within}}}$$

En pratique on veut R < 1.01

Détails en TD

1. Pourquoi Monte-Carlo ? (Exemple de modèle hiérarchique)
2. Introduction à la méthode Monte-Carlo (historique, PRNG)
3. Algorithmes de simulation i.i.d (PRNG, transformation, rejet)
4. Méthodes MCMC (Gibbs, Metropolis)
5. Diagonstics de convergence MCMC
6. Méthodes MCMC avancées (Langevin, HMC, NUTS)



$\widehat{R}$ = 1.005

Converging