Baysian Analysis

(1) Frequentist vs Baysian

	Frequentist	Baysian
θ	solution to an equation	random variable
p-val	probability of observing under H_0	probability of $\theta \in R$ Bayse Factor
C.I.	probability of $\theta \in I$ if repeated	probability of $\theta \in I$ Credible interval
principle	Likelihood (MLE)	posterior (MAP)

Def) $$p(\theta|x) = \frac{p(\theta, x)}{p(x)}$$

$$= \frac{p(\theta)p(x|\theta)}{p(x)}$$

$$\propto \underbrace{p(\theta)}_{\text{prior}} \underbrace{p(x|\theta)}_{\text{likelihood}}$$

* how to choose prior?

Def) A is a conjugate of B

$\Leftrightarrow$ $p(\theta|x)$ has the same distribution as $p(\theta)$

Example: $\mathcal{B}$ is a conjugate of Bin

MCMC

(1) Monte Carlo integration

$$E_{\theta \sim p(\theta|x)} f(\theta) \approx \frac{1}{M} \sum_{i=1}^{M} f(\theta_i)$$

θ_i is sampled from p

(2) Markov chain

Def) A Markov chain is a stochastic process s.t. $p(\theta_n \mid \theta_1, \cdots, \theta_{n-1})$

$= p(\theta_n \mid \theta_{n-1})$

Aim: Sample from a complex $p(\theta|x)$

(i) Initialize θ_0 s.t. $p(\theta_0) > 0$

(ii) Create a new sample $\theta_n \sim$

$f(\theta_n \mid \theta_{n-1}, x)$ where y is data and f is a transition density

(iii) Repeat!

- M-H Algorithm

$$\theta_n = \begin{cases} \theta^* & \alpha \text{ probability} \\ \theta_{n-1} & 1-\alpha \text{ probability} \end{cases}$$

where $\theta^* = \theta_{n-1} + Q$ (proposal)

$$\alpha = \min\left(1, \frac{p(\theta^* | x)}{p(\theta_{n-1} | x)} \cdot \frac{Q(\theta_{n-1} | \theta^*)}{Q(\theta^* | \theta_{n-1})}\right)$$

- Gibbs Algorithm

Let $\theta_0 = (\theta_0^1, \cdots, \theta_0^M)$

Update θ dimensionwisely

$$\theta_n^m \sim f(\theta_n^m \mid \theta_n^1, \cdots, \theta_n^{m-1}, \theta_{n-1}^{m+1}, \cdots, x, \theta_{n-1}^M)$$

<Remark> Burn-in Concept

<Remark 2> MLE vs MCMC

$\downarrow$

Sensitivity initial state vs model structure (or proposal)

Diagnostics of MC

* suggest using centered data

** Multiple chain idea (with various initial points)

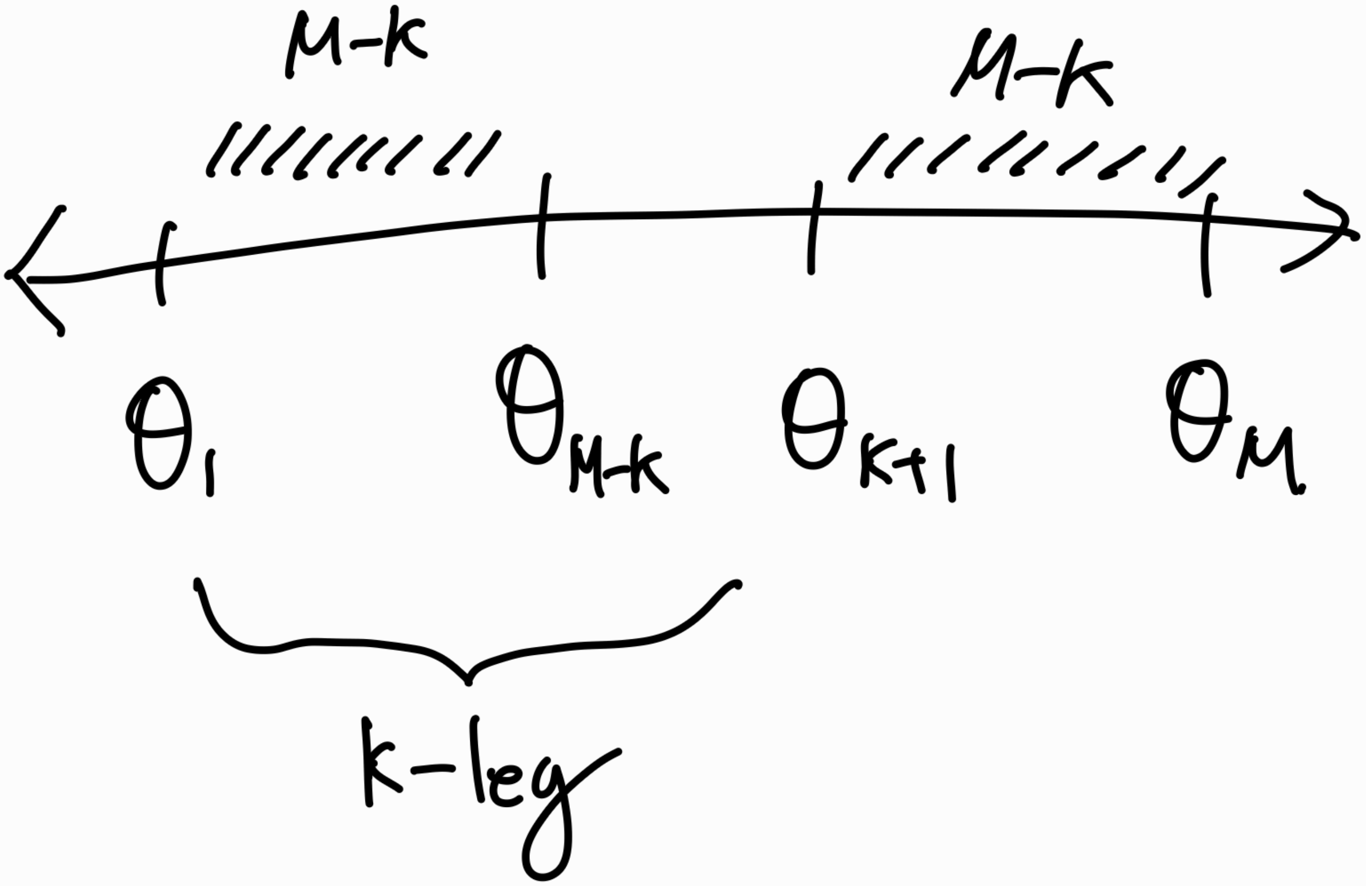
Def) Autocorrelation for a chain

$\Leftrightarrow$ For $\theta_1, \cdots, \theta_M$,

$$\rho(k) \equiv \frac{1}{\hat{\sigma}_\theta^2 (M-k)} \sum_{i=k+1}^{M} (\theta_i - \bar{\theta})(\theta_{i-k} - \bar{\theta})$$

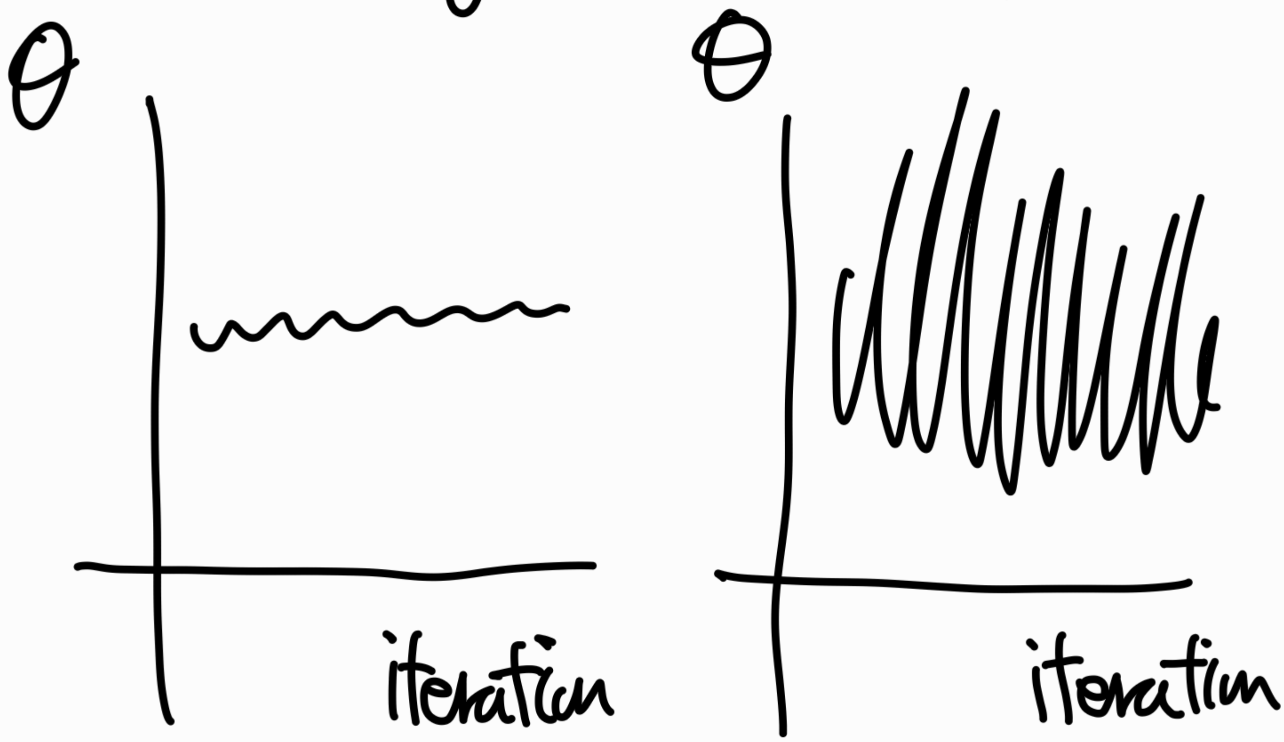
k is leg

sample variance

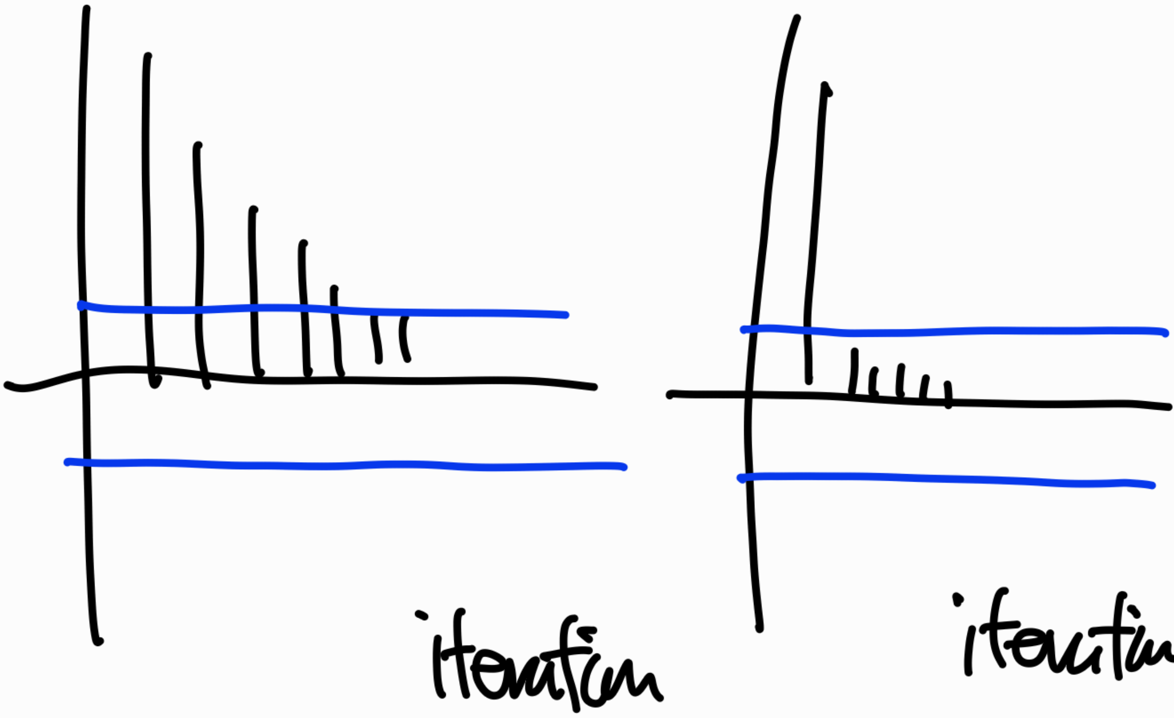


Example

No thinning   thinning = 3



ACF



confidence interval

$$\pm 1.96 \frac{1}{\sqrt{M-k}}$$