

Markov Chain Monte Carlo

Theory and Practical applications

Chapter 4: Geometric ergodicity and CLT

R. Douc and S. Schechtman

Télécom SudParis, Institut Polytechnique de Paris
randal.douc@telecom-sudparis.eu



- 1 Introduction
- 2 Coupling and total variation
- 3 Geometric ergodicity
- 4 Central Limit theorem

- 1 Introduction
- 2 Coupling and total variation
- 3 Geometric ergodicity
- 4 Central Limit theorem

Geometric ergodicity means that there exists constants $C > 0$ and $\varrho \in (0, 1)$ such that for all $n \in \mathbb{N}$,

$$\|\mu P^n - \pi\|_{\text{TV}} \leq C \varrho^n$$

where $\|\cdot\|_{\text{TV}}$ is the **total variation norm** (to be defined later) between two measures.

- 1 μP^n is the law of X_n starting from $X_0 \sim \mu$
- 2 π is the law of X_n starting from $X_0 \sim \pi$
- 3 **Geometric ergodicity for Markov chains** should not be confused with the notion of **ergodic dynamical systems**

CLT means that

$$n^{-1/2} \sum_{k=0}^{n-1} \{h(X_k) - \pi(h)\} \xrightarrow{\mathcal{L}_{\mathbb{P}_\pi}} \mathcal{N}(0, \sigma_\pi^2(h))$$

where h belongs to some class of functions and σ_π should be explicit.

Geometric ergodicity means that there exists constants $C > 0$ and $\varrho \in (0, 1)$ such that for all $n \in \mathbb{N}$,

$$\|\mu P^n - \pi\|_{\text{TV}} \leq C \varrho^n$$

where $\|\cdot\|_{\text{TV}}$ is the **total variation norm** (to be defined later) between two measures.

- 1 μP^n is the law of X_n starting from $X_0 \sim \mu$
- 2 π is the law of X_n starting from $X_0 \sim \pi$
- 3 **Geometric ergodicity for Markov chains** should not be confused with the notion of **ergodic dynamical systems**

CLT means that

$$n^{-1/2} \sum_{k=0}^{n-1} \{h(X_k) - \pi(h)\} \xrightarrow{\mathcal{L}_{\mathbb{P}, \pi}} \mathcal{N}(0, \sigma_{\pi}^2(h))$$

where h belongs to some class of functions and σ_{π} should be explicit.

Geometric ergodicity means that there exists constants $C > 0$ and $\varrho \in (0, 1)$ such that for all $n \in \mathbb{N}$,

$$\|\mu P^n - \pi\|_{\text{TV}} \leq C \varrho^n$$

where $\|\cdot\|_{\text{TV}}$ is the **total variation norm** (to be defined later) between two measures.

- 1 μP^n is the law of X_n starting from $X_0 \sim \mu$
- 2 π is the law of X_n starting from $X_0 \sim \pi$
- 3 **Geometric ergodicity for Markov chains** should not be confused with the notion of **ergodic dynamical systems**

CLT means that

$$n^{-1/2} \sum_{k=0}^{n-1} \{h(X_k) - \pi(h)\} \xrightarrow{\mathcal{L}_{\mathbb{P}_\pi}} \mathcal{N}(0, \sigma_\pi^2(h))$$

where h belongs to some class of functions and σ_π should be explicit.

Geometric ergodicity means that there exists constants $C > 0$ and $\varrho \in (0, 1)$ such that for all $n \in \mathbb{N}$,

$$\|\mu P^n - \pi\|_{\text{TV}} \leq C \varrho^n$$

where $\|\cdot\|_{\text{TV}}$ is the **total variation norm** (to be defined later) between two measures.

- 1 μP^n is the law of X_n starting from $X_0 \sim \mu$
- 2 π is the law of X_n starting from $X_0 \sim \pi$
- 3 **Geometric ergodicity for Markov chains** should not be confused with the notion of **ergodic dynamical systems**

CLT means that

$$n^{-1/2} \sum_{k=0}^{n-1} \{h(X_k) - \pi(h)\} \xrightarrow{\mathcal{L}_{\mathbb{P}_\pi}} \mathcal{N}(0, \sigma_\pi^2(h))$$

where h belongs to some class of functions and σ_π should be explicit.

Geometric ergodicity means that there exists constants $C > 0$ and $\varrho \in (0, 1)$ such that for all $n \in \mathbb{N}$,

$$\|\mu P^n - \pi\|_{\text{TV}} \leq C \varrho^n$$

where $\|\cdot\|_{\text{TV}}$ is the **total variation norm** (to be defined later) between two measures.

- 1 μP^n is the law of X_n starting from $X_0 \sim \mu$
- 2 π is the law of X_n starting from $X_0 \sim \pi$
- 3 **Geometric ergodicity for Markov chains** should not be confused with the notion of **ergodic dynamical systems**

CLT means that

$$n^{-1/2} \sum_{k=0}^{n-1} \{h(X_k) - \pi(h)\} \xrightarrow{\mathcal{L}_{\mathbb{P}_\pi}} \mathcal{N}(0, \sigma_\pi^2(h))$$

where h belongs to some class of functions and σ_π should be explicit.

- 1 Introduction
- 2 Coupling and total variation**
- 3 Geometric ergodicity
- 4 Central Limit theorem

Definition

Let (X, \mathcal{X}) be a measurable space and let ν, μ be two probability measures $\mu, \nu \in M_1(X)$. We define $\mathcal{C}(\mu, \nu)$, the **coupling set** associated to (μ, ν) as follows

$$\mathcal{C}(\mu, \nu) = \{ \gamma \in M_1(X^2) : \gamma(\cdot \times X) = \mu(\cdot), \gamma(X \times \cdot) = \nu(\cdot) \}$$

Any $\gamma \in \mathcal{C}(\mu, \nu)$ is called a **coupling** of (μ, ν) .

- 1 In words, γ is a **coupling of (μ, ν)** if the following property holds: if $(X, Y) \sim \gamma$, then we have the **marginal conditions**: $X \sim \mu$ and $Y \sim \nu$.
- 2 **Example:** The law of (X, X) where $X \sim \mu$ is a coupling of (μ, μ) . Other example if $X \sim \mu$ and $Y \sim \mu$ and $X \perp Y$, then, the law of (X, Y) is a coupling of (μ, μ) .

Definition

Let (X, \mathcal{X}) be a measurable space and let ν, μ be two probability measures $\mu, \nu \in M_1(X)$. We define $\mathcal{C}(\mu, \nu)$, the **coupling set** associated to (μ, ν) as follows

$$\mathcal{C}(\mu, \nu) = \{ \gamma \in M_1(X^2) : \gamma(\cdot \times X) = \mu(\cdot), \gamma(X \times \cdot) = \nu(\cdot) \}$$

Any $\gamma \in \mathcal{C}(\mu, \nu)$ is called a **coupling** of (μ, ν) .

- 1** In words, γ is a **coupling of (μ, ν)** if the following property holds: if $(X, Y) \sim \gamma$, then we have the **marginal conditions**: $X \sim \mu$ and $Y \sim \nu$.
- 2** **Example:** The law of (X, X) where $X \sim \mu$ is a coupling of (μ, μ) . Other example if $X \sim \mu$ and $Y \sim \mu$ and $X \perp Y$, then, the law of (X, Y) is a coupling of (μ, μ) .

Definition

Let (X, \mathcal{X}) be a measurable space and let ν, μ be two probability measures $\mu, \nu \in M_1(X)$. We define $\mathcal{C}(\mu, \nu)$, the **coupling set** associated to (μ, ν) as follows

$$\mathcal{C}(\mu, \nu) = \{ \gamma \in M_1(X^2) : \gamma(\cdot \times X) = \mu(\cdot), \gamma(X \times \cdot) = \nu(\cdot) \}$$

Any $\gamma \in \mathcal{C}(\mu, \nu)$ is called a **coupling** of (μ, ν) .

- 1** In words, γ is a **coupling of (μ, ν)** if the following property holds: if $(X, Y) \sim \gamma$, then we have the **marginal conditions**: $X \sim \mu$ and $Y \sim \nu$.
- 2** **Example:** The law of (X, X) where $X \sim \mu$ is a coupling of (μ, μ) . Other example if $X \sim \mu$ and $Y \sim \mu$ and $X \perp Y$, then, the law of (X, Y) is a coupling of (μ, μ) .

Definition

Let (X, \mathcal{X}) be a measurable space and let ν, μ be two probability measures $\mu, \nu \in M_1(X)$. Then the **total variation norm** between μ and ν noted $\|\mu - \nu\|_{TV}$, is defined by

$$\|\mu - \nu\|_{TV} = 2 \sup \{ |\mu(f) - \nu(f)| : f \in F(X), 0 \leq f \leq 1 \} \quad (1)$$

$$= \int |\varphi_0 - \varphi_1|(x) \zeta(dx) \quad (2)$$

$$= 2 \inf \{ \mathbb{P}(X \neq Y) : (X, Y) \sim \gamma \text{ where } \gamma \in \mathcal{C}(\mu, \nu) \} \quad (3)$$

where $\mu(dx) = \varphi_0(x)\zeta(dx)$ and $\nu(dx) = \varphi_1(x)\zeta(dx)$.

Proof is given in the lecture notes

- 1 Introduction
- 2 Coupling and total variation
- 3 Geometric ergodicity**
- 4 Central Limit theorem

Assumption A1

[**Minorizing condition**] for all $d > 0$, there exists $\epsilon_d > 0$ and a probability measure ν_d such that

$$\forall x \in C_d := \{V \leq d\}, \quad P(x, \cdot) \geq \epsilon_d \nu_d(\cdot) \quad (4)$$

Assumption A2

[**Drift condition**] there exists a constants $(\lambda, b) \in (0, 1) \times \mathbb{R}^+$ such that for all $x \in X$,

$$PV(x) \leq \lambda V(x) + b$$

Assumption A1

[**Minorizing condition**] for all $d > 0$, there exists $\epsilon_d > 0$ and a probability measure ν_d such that

$$\forall x \in C_d := \{V \leq d\}, \quad P(x, \cdot) \geq \epsilon_d \nu_d(\cdot) \quad (4)$$

Assumption A2

[**Drift condition**] there exists a constants $(\lambda, b) \in (0, 1) \times \mathbb{R}^+$ such that for all $x \in X$,

$$PV(x) \leq \lambda V(x) + b$$

Theorem

(Forgetting of the initialization) Assume (A1) and (A2) for some measurable function $V \geq 1$. Then, there exists a constant $\varrho \in (0, 1)$ such that for all $x, x' \in X$ and all $n \in \mathbb{N}$,

$$\|P^n(x, \cdot) - P^n(x', \cdot)\|_{\text{TV}} \leq \varrho^n [V(x) + V(x')].$$

Proof is hard. It is given in the lecture notes.

Theorem

(Forgetting of the initialization) Assume (A1) and (A2) for some measurable function $V \geq 1$. Then, there exists a constant $\rho \in (0, 1)$ such that for all $x, x' \in X$ and all $n \in \mathbb{N}$,

$$\|P^n(x, \cdot) - P^n(x', \cdot)\|_{\text{TV}} \leq \rho^n [V(x) + V(x')].$$

Proof is hard. It is given in the lecture notes.

Corollary

(Geometric ergodicity) Assume that (A1) and (A2) hold for some measurable function $V \geq 1$. Then, the Markov kernel P admits a *unique invariant probability measure* π . Moreover, $\pi(V) < \infty$ and there exists constants $(\varrho, \alpha) \in (0, 1) \times \mathbb{R}^+$ such that for all $\mu \in M_1(X)$ and all $n \in \mathbb{N}$,

$$\|\mu P^n - \pi\|_{TV} \leq \alpha \varrho^n \mu(V).$$

Proof should be done on the blackboard

Corollary

(Geometric ergodicity) Assume that (A1) and (A2) hold for some measurable function $V \geq 1$. Then, the Markov kernel P admits a *unique invariant probability measure* π . Moreover, $\pi(V) < \infty$ and there exists constants $(\varrho, \alpha) \in (0, 1) \times \mathbb{R}^+$ such that for all $\mu \in M_1(X)$ and all $n \in \mathbb{N}$,

$$\|\mu P^n - \pi\|_{\text{TV}} \leq \alpha \varrho^n \mu(V).$$

Proof should be done on the blackboard

Corollary

(Geometric ergodicity) Assume that (A1) and (A2) hold for some measurable function $V \geq 1$. Then, the Markov kernel P admits a *unique invariant probability measure* π . Moreover, $\pi(V) < \infty$ and there exists constants $(\varrho, \alpha) \in (0, 1) \times \mathbb{R}^+$ such that for all $\mu \in M_1(X)$ and all $n \in \mathbb{N}$,

$$\|\mu P^n - \pi\|_{TV} \leq \alpha \varrho^n \mu(V).$$

Proof should be done on the blackboard

- 1 Introduction
- 2 Coupling and total variation
- 3 Geometric ergodicity
- 4 Central Limit theorem**
 - Recap on martingales
 - The Poisson equation
 - Central limit theorems

Let $(M_n)_{n \in \mathbb{N}}$ be a sequence of random variables on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $(\mathcal{F}_n)_{n \in \mathbb{N}}$ be a filtration (ie for all $n \in \mathbb{N}$, $\mathcal{F}_n \subset \mathcal{F}_{n+1} \subset \mathcal{F}$).

Definition

We say that $(M_n)_{n \in \mathbb{N}}$ is a (\mathcal{F}_n) -martingale if for all $n \in \mathbb{N}$, M_n is integrable and for all $n \geq 1$,

$$\mathbb{E}[M_n | \mathcal{F}_{n-1}] = M_{n-1}$$

The *increment process* of the martingale is by definition $(M_{n+1} - M_n)_{n \in \mathbb{N}}$.

Theorem

If a sequence $(M_n)_{n \in \mathbb{N}}$ is a (\mathcal{F}_n) -martingale with stationary and square integrable increments, then

$$n^{-1/2} M_n \xrightarrow{\mathcal{L}_{\mathbb{P}}} \mathcal{N}(0, \mathbb{E}[(M_1 - M_0)^2])$$

Let $(M_n)_{n \in \mathbb{N}}$ be a sequence of random variables on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $(\mathcal{F}_n)_{n \in \mathbb{N}}$ be a filtration (ie for all $n \in \mathbb{N}$, $\mathcal{F}_n \subset \mathcal{F}_{n+1} \subset \mathcal{F}$).

Definition

We say that $(M_n)_{n \in \mathbb{N}}$ is a (\mathcal{F}_n) -martingale if for all $n \in \mathbb{N}$, M_n is integrable and for all $n \geq 1$,

$$\mathbb{E}[M_n | \mathcal{F}_{n-1}] = M_{n-1}$$

The *increment process* of the martingale is by definition $(M_{n+1} - M_n)_{n \in \mathbb{N}}$.

Theorem

If a sequence $(M_n)_{n \in \mathbb{N}}$ is a (\mathcal{F}_n) -martingale with stationary and square integrable increments, then

$$n^{-1/2} M_n \xrightarrow{\mathcal{L}_{\mathbb{P}}} \mathcal{N}(0, \mathbb{E}[(M_1 - M_0)^2])$$

Let $(M_n)_{n \in \mathbb{N}}$ be a sequence of random variables on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $(\mathcal{F}_n)_{n \in \mathbb{N}}$ be a filtration (ie for all $n \in \mathbb{N}$, $\mathcal{F}_n \subset \mathcal{F}_{n+1} \subset \mathcal{F}$).

Definition

We say that $(M_n)_{n \in \mathbb{N}}$ is a (\mathcal{F}_n) -martingale if for all $n \in \mathbb{N}$, M_n is integrable and for all $n \geq 1$,

$$\mathbb{E}[M_n | \mathcal{F}_{n-1}] = M_{n-1}$$

The *increment process* of the martingale is by definition $(M_{n+1} - M_n)_{n \in \mathbb{N}}$.

Theorem

If a sequence $(M_n)_{n \in \mathbb{N}}$ is a (\mathcal{F}_n) -martingale with stationary and square integrable increments, then

$$n^{-1/2} M_n \xrightarrow{\mathcal{L}_{\mathbb{P}}} \mathcal{N}(0, \mathbb{E}[(M_1 - M_0)^2])$$

Definition

For a given measurable function h such that $\pi|h| < \infty$, the **Poisson equation** is defined by

$$\hat{h} - P\hat{h} = h - \pi(h) \quad (5)$$

A solution to the Poisson equation is a function \hat{h} for which (5) holds provided that $P|\hat{h}|(x) < \infty$ for all $x \in X$.

Link between Poisson equations and Martingales

Define

$$\begin{aligned} S_n(h) &= \sum_{k=0}^{n-1} \{h(X_k) - \pi(h)\} \\ &= M_n(\hat{h}) + \hat{h}(X_0) - \hat{h}(X_n) \end{aligned}$$

where

$$M_n(\hat{h}) = \sum_{k=1}^n \left\{ \hat{h}(X_k) - P\hat{h}(X_{k-1}) \right\} \quad (6)$$

Note that $\{M_n(\hat{h})\}_{n \in \mathbb{N}}$ is indeed a (\mathcal{F}_k) -martingale where $\mathcal{F}_k = \sigma(X_0, \dots, X_k)$. Indeed we have

$$\begin{aligned} \mathbb{E}[M_n(\hat{h}) | \mathcal{F}_{n-1}] - M_{n-1}(h) &= \mathbb{E}[\hat{h}(X_n) - P\hat{h}(X_{n-1}) | \mathcal{F}_{n-1}] \\ &= P\hat{h}(X_{n-1}) - P\hat{h}(X_{n-1}) = 0 \end{aligned}$$

Link between Poisson equations and Martingales

Define

$$\begin{aligned} S_n(h) &= \sum_{k=0}^{n-1} \{h(X_k) - \pi(h)\} \\ &= M_n(\hat{h}) + \hat{h}(X_0) - \hat{h}(X_n) \end{aligned}$$

where

$$M_n(\hat{h}) = \sum_{k=1}^n \left\{ \hat{h}(X_k) - P\hat{h}(X_{k-1}) \right\} \quad (6)$$

Note that $\{M_n(\hat{h})\}_{n \in \mathbb{N}}$ is indeed a (\mathcal{F}_k) -martingale where $\mathcal{F}_k = \sigma(X_0, \dots, X_k)$. Indeed we have

$$\begin{aligned} \mathbb{E}[M_n(\hat{h}) | \mathcal{F}_{n-1}] - M_{n-1}(h) &= \mathbb{E}[\hat{h}(X_n) - P\hat{h}(X_{n-1}) | \mathcal{F}_{n-1}] \\ &= P\hat{h}(X_{n-1}) - P\hat{h}(X_{n-1}) = 0 \end{aligned}$$

Link between Poisson equations and Martingales

Define

$$\begin{aligned} S_n(h) &= \sum_{k=0}^{n-1} \{h(X_k) - \pi(h)\} \\ &= M_n(\hat{h}) + \hat{h}(X_0) - \hat{h}(X_n) \end{aligned}$$

where

$$M_n(\hat{h}) = \sum_{k=1}^n \left\{ \hat{h}(X_k) - P\hat{h}(X_{k-1}) \right\} \quad (6)$$

Note that $\{M_n(\hat{h})\}_{n \in \mathbb{N}}$ is indeed a (\mathcal{F}_k) -martingale where $\mathcal{F}_k = \sigma(X_0, \dots, X_k)$. Indeed we have

$$\begin{aligned} \mathbb{E}[M_n(\hat{h}) | \mathcal{F}_{n-1}] - M_{n-1}(h) &= \mathbb{E}[\hat{h}(X_n) - P\hat{h}(X_{n-1}) | \mathcal{F}_{n-1}] \\ &= P\hat{h}(X_{n-1}) - P\hat{h}(X_{n-1}) = 0 \end{aligned}$$

Link between Poisson equations and Martingales

Define

$$\begin{aligned} S_n(h) &= \sum_{k=0}^{n-1} \{h(X_k) - \pi(h)\} \\ &= M_n(\hat{h}) + \hat{h}(X_0) - \hat{h}(X_n) \end{aligned}$$

where

$$M_n(\hat{h}) = \sum_{k=1}^n \left\{ \hat{h}(X_k) - P\hat{h}(X_{k-1}) \right\} \quad (6)$$

Note that $\left\{ M_n(\hat{h}) \right\}_{n \in \mathbb{N}}$ is indeed a (\mathcal{F}_k) -martingale where $\mathcal{F}_k = \sigma(X_0, \dots, X_k)$. Indeed we have

$$\begin{aligned} \mathbb{E}[M_n(\hat{h}) | \mathcal{F}_{n-1}] - M_{n-1}(h) &= \mathbb{E}[\hat{h}(X_n) - P\hat{h}(X_{n-1}) | \mathcal{F}_{n-1}] \\ &= P\hat{h}(X_{n-1}) - P\hat{h}(X_{n-1}) = 0 \end{aligned}$$

Link between Poisson equations and Martingales

Define

$$\begin{aligned} S_n(h) &= \sum_{k=0}^{n-1} \{h(X_k) - \pi(h)\} \\ &= M_n(\hat{h}) + \hat{h}(X_0) - \hat{h}(X_n) \end{aligned}$$

where

$$M_n(\hat{h}) = \sum_{k=1}^n \left\{ \hat{h}(X_k) - P\hat{h}(X_{k-1}) \right\} \quad (6)$$

Note that $\{M_n(\hat{h})\}_{n \in \mathbb{N}}$ is indeed a (\mathcal{F}_k) -martingale where $\mathcal{F}_k = \sigma(X_0, \dots, X_k)$. Indeed we have

$$\begin{aligned} \mathbb{E}[M_n(\hat{h}) | \mathcal{F}_{n-1}] - M_{n-1}(h) &= \mathbb{E}[\hat{h}(X_n) - P\hat{h}(X_{n-1}) | \mathcal{F}_{n-1}] \\ &= P\hat{h}(X_{n-1}) - P\hat{h}(X_{n-1}) = 0 \end{aligned}$$

Theorem

Assume that (A1) and (A2) hold for some measurable function $V \geq 1$. Then, for any function h such that $|h| \leq V$, the function

$$\hat{h} = \sum_{n=0}^{\infty} \{P^n h - \pi(h)\} \quad (7)$$

is well-defined. Moreover, \hat{h} is a **solution of the Poisson equation** associated to h and there exists a constant γ such that for all $x \in X$,

$$|\hat{h}(x)| \leq \gamma V(x)$$

Proof should be done on the blackboard

Theorem

(CLT with Poisson assumption) Let P be a Markov kernel with a unique invariant probability measure π . Let $h \in L^2(\pi)$. Assume that there exists a solution $\hat{h} \in L^2(\pi)$ to the Poisson equation $\hat{h} - P\hat{h} = h$. Then

$$n^{-1/2} \sum_{k=0}^{n-1} \{h(X_k) - \pi(h)\} \xrightarrow{\mathcal{L}_{\mathbb{P}_\pi}} \mathcal{N}(0, \sigma_\pi^2(h)),$$

where

$$\sigma_\pi^2(h) = \mathbb{E}_\pi[\{\hat{h}(X_1) - P\hat{h}(X_0)\}^2] \quad (8)$$

Proof should be done on the blackboard

Theorem

(CLT with Poisson assumption) Let P be a Markov kernel with a unique invariant probability measure π . Let $h \in L^2(\pi)$. Assume that there exists a solution $\hat{h} \in L^2(\pi)$ to the Poisson equation $\hat{h} - P\hat{h} = h$. Then

$$n^{-1/2} \sum_{k=0}^{n-1} \{h(X_k) - \pi(h)\} \xrightarrow{\mathcal{L}_{\mathbb{P}_\pi}} \mathcal{N}(0, \sigma_\pi^2(h)),$$

where

$$\sigma_\pi^2(h) = \mathbb{E}_\pi[\{\hat{h}(X_1) - P\hat{h}(X_0)\}^2] \quad (8)$$

Proof should be done on the blackboard

Theorem

(CLT with A1-A2 assumptions) Assume that (A1 and (A2) hold for some function V . Then, for all measurable functions h such that $|h|^2 \leq V$,

$$n^{-1/2} \sum_{k=0}^{n-1} \{h(X_k) - \pi(h)\} \stackrel{\mathcal{L}_{\mathbb{P}_\pi}}{\Rightarrow} \mathcal{N}(0, \sigma_\pi^2(h)),$$

where

$$\sigma_\pi^2(h) = \mathbb{E}_\pi[\{\hat{h}(X_1) - P\hat{h}(X_0)\}^2] \quad (9)$$

and \hat{h} is defined as in (7).

Proof should be done on the blackboard

Theorem

(CLT with A1-A2 assumptions) Assume that (A1 and (A2) hold for some function V . Then, for all measurable functions h such that $|h|^2 \leq V$,

$$n^{-1/2} \sum_{k=0}^{n-1} \{h(X_k) - \pi(h)\} \stackrel{\mathcal{L}_{\mathbb{P}_\pi}}{\Rightarrow} \mathcal{N}(0, \sigma_\pi^2(h)),$$

where

$$\sigma_\pi^2(h) = \mathbb{E}_\pi[\{\hat{h}(X_1) - P\hat{h}(X_0)\}^2] \quad (9)$$

and \hat{h} is defined as in (7).

Proof should be done on the blackboard