2018-03-17

Statistical Inference Principles

StatInfPr 18-1!

Re-exam läsår 2017/18, LP 4

Examiner: Annika Lang

Problem 1 $\boxed{\Sigma\, 26}$

Let the pmf

$$f(x\mid\theta) = \theta^x (1-\theta)^x, \quad x \in \{0,1\}, \quad \theta \in [0, \tfrac{1}{2}]$$

be given, i.e. the corresponding RV is Bernoulli with parameter θ.

③ (a) A family of pdfs or pmfs is called an *exponential family* if it has a representation

$$f(x\mid\theta) = h(x)\, c(\theta) \exp\left(\sum_{i=1}^{k} w_i(\theta)\, t_i(x)\right) \qquad (1)$$

with

- $h : \mathbb{R} \to \mathbb{R}_+$, i.e., $h(x) \geq 0 \ \forall x \in \mathbb{R}$
- $t_1, \dots, t_k : \mathbb{R} \to \mathbb{R}$
- $c : \mathbb{R}^d \to \mathbb{R}_+$, i.e., $c(\theta) \geq 0 \ \forall \theta \in \mathbb{R}^d$
- $w_1, \dots, w_k : \mathbb{R}^d \to \mathbb{R}$.

Claim: $\{f(\cdot\mid\theta),\ \theta \in [0, \tfrac{1}{2}]\}$ is an exponential family.

Proof:

Observe that

$$\begin{aligned} f(x\mid\theta) &= \theta^x (1-\theta)^{1-x} \\ &= (1-\theta) \cdot \theta^x (1-\theta)^{-x} \\ &= (1-\theta) \cdot \left(\frac{\theta}{1-\theta}\right)^x \\ &= \mathbb{1}_{\{0,1\}}(x)\, (1-\theta) \exp\left(\ln\left(\frac{\theta}{1-\theta}\right)^x\right) \\ &= \underbrace{\mathbb{1}_{\{0,1\}}(x)}_{h(x) \geq 0\ (\frac{1}{2})}\ \underbrace{(1-\theta)}_{c(\theta) \geq 0\ (\frac{1}{2})} \exp\Big(\underbrace{\ln\frac{\theta}{1-\theta}}_{w_1(\theta)\ (\frac{1}{2})} \cdot \underbrace{x}_{t_1(x)\ (\frac{1}{2})}\Big), \end{aligned}$$

i.e., f can be written in the required representation with $k=1$.

4)

(b) A statistic $T(\underline{X})$ of a random sample $\underline{X}$ is a <u>sufficient statistic</u> for the parameter θ if the conditional distribution of $\underline{X}$ given $T(\underline{X})$ does not depend on θ. (1/2)

The Factorization Theorem states that $T(\underline{X})$ is sufficient statistic for θ $\iff$ (1/2)

$$\exists\, g(\cdot|\theta) \text{ and } h: \quad \forall x \in \mathcal{X}\ \forall \theta \in \Theta: \underbrace{f(x|\theta)}_{\text{joint pdf/pmf of } \underline{X}} = g(T(x)|\theta)\, h(x)$$

Let us consider the exponential family representation:

$$f(x|\theta) = \mathbb{1}_{\{0,1\}}(x)\,(1-\theta)\exp\left(\ln\tfrac{\theta}{1-\theta}\, x\right)$$

Then the joint pmf of a RS of size n is

$$\prod_{i=1}^{n} f(x_i|\theta) \overset{(1/2)}{=} \prod_{i=1}^{n}\left(\mathbb{1}_{\{0,1\}}(x_i)\,(1-\theta)\exp\left(\ln\tfrac{\theta}{1-\theta}\, x_i\right)\right)$$

$$\overset{(1/2)}{=} \mathbb{1}_{\{0,1\}^n}(x)\,(1-\theta)^n \exp\left(\ln\tfrac{\theta}{1-\theta}\sum_{i=1}^{n} x_i\right)$$

$$= \mathbb{1}_{\{0,1\}^n}(x)\, \underbrace{(1-\theta)^n \left(\tfrac{\theta}{1-\theta}\right)^{\sum_{i=1}^{n} x_i}}_{\|}$$

$$= \overset{(1/2)}{h(x)} \cdot \overset{(1/2)}{g}\left(\sum_{i=1}^{n} x_i \,\Big|\, \theta\right)$$

$\Rightarrow \sum_{i=1}^{n} x_i$ is a sufficient statistic for θ. (1)

4)

(c) A sufficient statistic $T(\underline{X})$ for θ of the RS $\underline{X}$ is called a <u>minimal sufficient statistic</u> if for any sufficient statistic $T'(\underline{X})$, T is a function of T', i.e., (1/2)

$$T'(x) = T'(y) \Rightarrow T(x) = T(y)$$

Theorem (Lehmann, Scheffé)

$f(\cdot|\theta)$ pmf or pdf of RS $\underline{X}$

Assume:

$$\exists T: \forall x, y \in \mathcal{X}: \frac{f(x|\theta)}{f(y|\theta)} = \text{const as fction of } \theta$$
$$\Leftrightarrow T(x) = T(y)$$

(1/2) -1

Then $T(\underline{X})$ is minimal sufficient statistic for θ.

Let us continue the computations of (b). We obtain

$$\frac{\prod_{i=1}^{n} f(x_i|\theta)}{\prod_{j=1}^{n} f(y_j|\theta)} \overset{(1/2)}{=} \frac{\mathbb{1}_{\{0,1\}^n}(x)}{\underbrace{\mathbb{1}_{\{0,1\}^n}(y)}_{\text{assumed} \neq 0}} \frac{(1-\theta)^n}{(1-\theta)^n} \cdot \frac{\left(\frac{\theta}{1-\theta}\right)^{\sum_{i=1}^n x_i}}{\left(\frac{\theta}{1-\theta}\right)^{\sum_{j=1}^n y_j}}$$

$$\overset{(1/2)}{=} \mathbb{1}_{\{0,1\}^n}(x)\, \mathbb{1}_{\{0,1\}^n}(y) \cdot \left(\frac{\theta}{1-\theta}\right)^{\sum_{i=1}^n x_i - \sum_{j=1}^n y_j}$$

This is constant as a function of θ

$$\overset{(1)}{\Leftrightarrow} \sum_{i=1}^{n} x_i = \sum_{j=1}^{n} y_j \quad \left(\text{or } \sum_{i=1}^n x_i - \sum_{j=1}^n y_j = 0\right)$$

By Lehmann/Scheffé, $\sum_{i=1}^{n} x_i$ is also minimal sufficient statistic. (1)

(3) (d) Let $\underline{X} = (X_1, .., X_n)$ be a RS from a population with pmf

$$f(x|\theta) = \theta^x (1-\theta)^{1-x}, \quad x \in \{0,1\}, \ \theta \in [0, \tfrac{1}{2}]$$

The method of moments matches the first k moments, where k denotes the dimension of θ, i.e., $k = 1$ here. (1/2)

We compute

$$E(X_1) = 0 \cdot \theta^0 (1-\theta)^{1-0} + 1 \cdot \theta^1 (1-\theta)^{1-1} \quad (1/2)$$
$$= \theta \quad (1/2)$$

and $m_1 = \frac{1}{n} \sum_{i=1}^{n} X_i$ (1/2)

$$\Rightarrow \tilde{\theta} = \overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \quad (1)$$

6

(e) Let $f(\cdot\,|\theta)$ denote the joint pdf or pmf of the RS $\underline{X}$. Given that $\underline{X} = x$ is observed, the <u>likelihood function</u> L of θ is given by

$$L(\theta|x) := f(x|\theta).$$

The <u>maximum likelihood estimator</u> of θ based on a sample $\underline{X}$ is $\hat{\theta}(\underline{X})$ or $\hat{\theta}$ which maximizes $L(\theta|\underline{X})$. 1/2

Compute first $L(\theta|X)$.

$$L(\theta|x) = f(\theta|(x_1, \dots, x_n))$$
$$= \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i}$$
$$= \theta^{\sum_{i=1}^{n} x_i} (1-\theta)^{n-\sum_{i=1}^{n} x_i} \quad 1/2$$
$$\left[= (1-\theta)^n \left(\frac{\theta}{1-\theta}\right)^{\sum_{i=1}^{n} x_i}\right]$$

Since the log likelihood maximizes at the same values, differentiate this:

$$\frac{\partial}{\partial\theta} \ln L(\theta|x) \overset{1/2}{=} \frac{\partial}{\partial\theta}\left(\sum_{i=1}^{n} x_i \ln\theta + \left(n - \sum_{i=1}^{n} x_i\right)\ln(1-\theta)\right)$$
$$\overset{\theta \neq 0}{=} \frac{1}{\theta}\sum_{i=1}^{n} x_i + \frac{1}{1-\theta}(-1)\left(n - \sum_{i=1}^{n} x_i\right) \quad 1/2$$
$$= \theta^{-1}\sum_{i=1}^{n} x_i - (1-\theta)^{-1}\left(n - \sum_{i=1}^{n} x_i\right)$$
$$\overset{!}{=} 0 \quad 1/2$$

$$\Leftrightarrow \quad (1-\cancel{\theta})\sum_{i=1}^{n} x_i - \theta\left(n - \cancel{\sum_{i=1}^{n} x_i}\right) = \sum_{i=1}^{n} x_i - n\theta \overset{!}{=} 0$$

$$\Leftrightarrow \quad \theta = \frac{1}{n}\sum_{i=1}^{n} x_i \quad 1/2$$

and
$$\frac{\partial^2}{\partial\theta^2} \ln L(\theta|x) = -\theta^{-2}\sum_{i=1}^{n} x_i - (1-\theta)^{-2}\left(n - \sum_{i=1}^{n} x_i\right)$$
$$\hat{=} -n^2\left(\sum_{i=1}^{n} x_i\right)^{-1} - \frac{n - \sum_{i=1}^{n} x_i}{\left(1 - \frac{1}{n}\sum_{i=1}^{n} x_i\right)^2}$$
$$\hat{=} -n^2\left(\sum_{i=1}^{n} x_i\right)^{-1} - n^2\left(n - \sum_{i=1}^{n} x_i\right)^{-1}$$

$$= -n^2 \left(\frac{(1-\sum_{i=1}^n x_i) + \sum_{i=1}^n x_i}{(\sum_{i=1}^n x_i)(n-\sum_{i=1}^n x_i)} \right)$$

$$= -n^2 \left(\underbrace{\sum_{i=1}^n x_i}_{>0} \cdot \underbrace{(n-\sum_{i=1}^n x_i)}_{>0} \right)^{-1} \quad \text{if not equal } 0 \text{ or } 1$$

$$< 0$$

(1)

$\Rightarrow \quad \hat{\theta} := \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}$ is a maximum (1/2)

$\Rightarrow \quad \hat{\theta}$ is the MLE of θ if $\hat{\theta} \leq \frac{1}{2}$.

Since L is an increasing function in θ, it obtains its maximum at the upper bound, i.e, $\frac{1}{2}$ (1/2)

$\Rightarrow \quad \underline{\hat{\theta} = \min\{\bar{X}, \frac{1}{2}\}}$ is the MLE (1)

(2½) (f) The mean squared error is:

<u>Method of moments:</u>

$$MSE = \mathbb{E}((\tilde{\theta} - \theta)^2) \quad (1/2)$$

$$= \mathbb{E}((\tfrac{1}{n} \sum_{i=1}^n X_i - \theta)^2)$$

$$\overset{\text{unbiased est. of } \theta}{=} \mathbb{E}((\tfrac{1}{n} \sum_{i=1}^n X_i - \mathbb{E}(\tfrac{1}{n} \sum_{i=1}^n X_i))^2)$$

$$= Var(\tfrac{1}{n} \sum_{i=1}^n X_i) \quad (1/2)$$

$$(1/2) \overset{\text{iid, Bienaymé}}{=} \frac{1}{n^2} \sum_{i=1}^n Var(X_i)$$

$$= \frac{1}{n} Var(X_1)$$

$$= \frac{1}{n} \left(1^2 \theta^1 (1-\theta)^0 + 0 - \theta^2 \right) \quad (1/2)$$

$$= \frac{\theta(1-\theta)}{n} \quad (1/2)$$

(3½) (g) A family of pmfs $\{g(\cdot|\theta), \theta \in \Theta\}$, $\theta \in \mathbb{R}$, has a <u>monotone likelihood ratio</u> if

$\forall\, \theta_2 > \theta_1$: $\frac{g(\cdot|\theta_2)}{g(\cdot|\theta_1)}$ is a monotone (nondecreasing) function on

$\{ t \in \mathcal{T}, \ g(t|\theta_1) > 0 \text{ or } g(t|\theta_2) > 0 \}$, (1/2)

where $\frac{c}{0} := +\infty$ for $c > 0$.

It holds that

$$\frac{f(x|\theta_2)}{f(x|\theta_1)} \overset{(1/2)}{=} \frac{\cancel{\mathbb{1}_{\{0,1\}}(x)}\,(1-\theta_2)\,\exp\left(\ln\frac{\theta_2}{1-\theta_2}\cdot x\right)}{\cancel{\mathbb{1}_{\{0,1\}}(x)}\,(1-\theta_1)\,\exp\left(\ln\frac{\theta_1}{1-\theta_1}\cdot x\right)}$$

$$\overset{(1/2)}{=} \underbrace{\frac{1-\theta_2}{1-\theta_1}}_{>0 \ (1/2)} \exp\Big(\underbrace{\left(\ln\frac{\theta_2}{1-\theta_2} - \ln\frac{\theta_1}{1-\theta_1}\right)}_{\text{since increasing } >0 \ (1/2)} x\Big)$$

$\Rightarrow$ increasing function of x (1/2)

$\Rightarrow \{ f(\cdot|\theta), \theta \in [0, \frac{1}{2}] \}$ has a monotone likelihood ratio. (1/2)

Problem 2 $\boxed{\Sigma\, 13}$

(5 1/2)

(a) We recall from the first exam for a unimodal pdf f:

If the interval $[a,b]$ satisfies

(i) $\int_a^b f(x)\,dx = 1-\alpha$, (1)

(ii) $f(a) = f(b) > 0$,

(iii) $a \le x^* \le b$, where x^* is a mode of f,

then $[a,b]$ is a shortest among all intervals that satisfy (i).

To prove the claim we have to show that (i) – (iii) are satisfied, since f is unimodal by assumption (1/2)

(i) $$\int_a^b f(x)\,dx \overset{(1/2)}{=} \int_{-\infty}^{\infty} f(x)\,dx - \int_{-\infty}^{a} f(x)\,dx - \int_b^{\infty} f(x)\,dx$$

$$= 1 - \frac{\alpha}{2} - \frac{\alpha}{2}$$

$$= 1 - \alpha \quad (1/2)$$

(ii) Since f is symmetric, $\int_{-\infty}^{a} f(x)\,dx = \int_b^{+\infty} f(x)\,dx \Rightarrow f(a) = f(b)$

with $b = \mu + (\mu - a) = 2\mu - a$, where μ is the point (1)

of symmetry with $\int_{-\infty}^{\mu} f(x)\,dx = \frac{1}{2}$

$\Rightarrow f$ is increasing on $(-\infty, a]$ and since $\int_{-\infty}^{a} f(x)dx \geq 0$ (1/2)
this implies that $f(a) \geq 0$. (1/2)

(iii) The point of symmetry μ is a mode x^* and therefore (1/2)
$a \leq \mu \leq b$ (1/2) with $a = \mu = b$ for $\alpha = 1$.

Since all conditions are fullfilled, the given $[a,b]$ is a shortest interval.

3 (b) Let $f(x) = \mathbb{1}_{[0,1]}(x)$, then f is the pdf of the uniform distribution on $[0,1]$, which is symmetric with point of (1) symmetry $\mu = \frac{1}{2}$. Fix $\alpha \in (0, \frac{1}{2})$
By (a), a shortest interval is given by $[a,b]$ s.t.

(1) $$\int_{-\infty}^{a} f(x)\,dx = \int_{b}^{\infty} f(x)\,dx = \frac{\alpha}{2},$$

i.e, $a = \frac{\alpha}{2}$ and $b = 1 - \frac{\alpha}{2}$.

The interval $[a,b]$ has length $b - a = 1 - \frac{\alpha}{2} - \frac{\alpha}{2} = 1 - \alpha$.
Choose $a' = 0$ and $b' = 1-\alpha$, then

(1) $$\int_{a'}^{b'} f(x)\,dx = \int_{0}^{1-\alpha} \mathbb{1}_{[0,1]}(x)\,dx = 1-\alpha$$

with length $b' - a' = 1 - \alpha - 0 = 1 - \alpha = b - a$,
i.e., $[a', b']$ has the same length and integrates to $1-\alpha$
but $[a,b] \neq [a',b']$.

4½ (c) $(X_i, i=1,\dots,n)$ RS, $N(\mu, \sigma^2)$ distributed

- The sample mean $\overline{X}$ is $N(\mu, \frac{\sigma^2}{n})$ distributed (1/2)
with pdf

$$f(x) = \frac{1}{\sqrt{2\pi\frac{\sigma^2}{n}}} e^{-\frac{(x-\mu)^2}{2\frac{\sigma^2}{n}}},$$

which is symmetric wrt $\mu =: x^*$

Furthermore

$$Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1) \quad (1)$$

with corresponding density $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, which is symmetric around 0. (1/2)

$\Rightarrow \exists\ a < 0$ s.t. $\int_{-\infty}^{a} f(x)\,dx = \frac{\alpha}{2}$ (1/2)

and for $b = -a$ (1/2) $\int_{b}^{\infty} f(x)\,dx = \frac{\alpha}{2}$ (1/2)

$\overset{(a)}{\Rightarrow}$ $[a, -a] = [-b, b]$ is a shortest $1-\alpha$ confidence interval for 0. (1/2)

$\Rightarrow \{\mu \in \mathbb{R},\ -b \leq z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} \leq b\}$

$= \{\mu \in \mathbb{R},\ \bar{x} - b\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + b\frac{\sigma}{\sqrt{n}}\}$ (1/2)

is a shortest $1-\alpha$ confidence interval for $\mu = \mathbb{E}(X)$.

Problem 3 $\boxed{\Sigma 16}$

Neyman-Pearson Lemma

Test $H_0: \theta = \theta_0$ vs. $H_1: \theta = \theta_1$

with pmfs $f(\cdot|\theta_i)$, $i = 0, 1$

and rejection region R that satisfies:

$\exists k \geq 0$,

$$x \in R \quad \text{if} \quad f(x|\theta_1) > k\, f(x|\theta_0) \qquad (8.3.1)$$

and

$$x \in R^c \quad \text{if} \quad f(x|\theta_1) < k\, f(x|\theta_0)$$

and

$$\alpha = P_{\theta_0}(X \in R) \qquad (8.3.2)$$

assume test if can be attained!

Then,

(a) (Sufficiency) Any test that satisfies (8.3.1) and (8.3.2) is a UMP level α test

(b) (Necessity) If there exists a test satisfying (8.3.1) and (8.3.2) with $k > 0$, then every UMP level α test is a size α test (satisfies (8.3.2)) and every UMP level α test satisfies (8.3.1) except perhaps on a set A satisfying $P_{\theta_0}(X \in A) = P_{\theta_1}(X \in A) = 0$

Proof: (similarly to the continuous case, i.e., $f(\cdot|\theta_i)$, $i=0,1$, pdfs)

$3\frac{1}{2}$ (a) Assume we are given a level α test with

- rejection region R $\frac{1}{2}$
- power function β $\frac{1}{2}$
- test function ϕ given by

$$\phi(x) = \begin{cases} 1 & \text{if } x \in R \\ 0 & \text{else} \end{cases} \qquad 1$$

satisfying (8.3.1) and (8.3.2)

$\overset{(8.3.2)}{\Rightarrow}$ test is a size α test and $\Theta_0 = \{\theta_0\}$, $\Theta_0^c = \{\theta_1\}$ 1

Furthermore assume any other level α test with

- rejection region R' } 1

- power function β'
- test function ϕ'

Then

$(\frac{1}{2})$ $$\underbrace{(\phi(x) - \phi'(x))}_{(\frac{1}{2})\begin{cases}\geq 0 & x \in R \\ \leq 0 & x \in R^c\end{cases}} \underbrace{(f(x|\theta_1) - k f(x|\theta_0))}_{(\frac{1}{2})\begin{cases}> 0 & x \in R \\ < 0 & x \in R^c\end{cases}\ (8.3.1)} \geq 0 \ (\tfrac{1}{2})$$

$$\Rightarrow 0 \overset{(\frac{1}{2})}{\leq} \sum_{x \in \mathcal{X}} (\phi(x) - \phi'(x))(f(x|\theta_1) - k f(x|\theta_0))$$

$$\overset{(\frac{1}{2})}{=} \sum_{x \in R} (f(x|\theta_1) - k f(x|\theta_0)) - \sum_{x \in R'} (f(x|\theta_1) - k f(x|\theta_0))$$

$$\overset{(\frac{1}{2})}{=} P_{\theta_1}(X \in R) - k \underbrace{P_{\theta_0}(X \in R)}_{= \alpha} - (P_{\theta_1}(X \in R') - k P_{\theta_0}(X \in R'))$$

$$\overset{(\frac{1}{2})}{=} \beta(\theta_1) - \beta'(\theta_1) - \underbrace{\underbrace{k}_{\geq 0}(\alpha - \underbrace{\beta'(\theta_0)}_{\leq \alpha})}_{\geq 0}$$

$$\overset{(\frac{1}{2})}{\leq} \beta(\theta_1) - \beta'(\theta_1)$$

$\Leftrightarrow (\frac{1}{2})$ $\beta'(\theta_1) \leq \beta(\theta_1)$, i.e., β is UMP test $(\frac{1}{2})$ since $\{\theta_1\} = \Theta_0^c$ and second test arb.

(b) Setup of (a) but ϕ' now test function of a UMP level α test and $k > 0$ (1)

$1\frac{1}{2}$

$$\Rightarrow (\tfrac{1}{2})\ \beta(\theta_1) = \beta'(\theta_1)$$

$$\overset{(a)}{\Rightarrow} (1)\ 0 \leq -\underbrace{k}_{> 0}(\alpha - \underbrace{\beta'(\theta_0)}_{\leq \alpha})$$

$\Rightarrow (\frac{1}{2})$ $\beta'(\theta_0) = \alpha$, i.e., size α test $(\frac{1}{2})$ that satisfies (8.3.2)

$$\Rightarrow (\tfrac{1}{2})\ 0 = \sum_{x \in \mathcal{X}} \overbrace{(\phi(x) - \phi'(x)) \underbrace{(f(x|\theta_1) - k f(x|\theta_0))}_{\neq 0}}^{\geq 0}$$

$$\Rightarrow (\tfrac{1}{2})\ (\phi(x) - \phi'(x))(f(x|\theta_1) - k f(x|\theta_0)) = 0 \quad \forall x \in \mathcal{X}$$

$$\Rightarrow (\tfrac{1}{2})\ f(x|\theta_1) = f(x|\theta_0) = 0 \text{ or } (\tfrac{1}{2})\ \phi(x) = \phi'(x) \quad \forall x \in \mathcal{X}$$

$\Rightarrow (\frac{1}{2})$ $R \setminus A = R' \setminus A$, where $A = \{x \in \mathcal{X},\ f(x|\theta_1) = f(x|\theta_0) = 0\}$

$\Rightarrow$ second test satisfies (8.3.1) (since test dependent on R) $(\frac{1}{2})$

□

(S)

StatInfPrinc 18-6

CHALMERSSKA HUSET

Problem 4 $\boxed{\Sigma 5}$

$\Sigma 5$

1 (a) A sequence of estimators $(W_n, n \in \mathbb{N})$ is a consistent sequence of estimators of the parameter θ if (1)

$$\forall \varepsilon > 0 : \forall \theta \in \Theta : \lim_{n \to +\infty} P_\theta(|W_n - \theta| < \varepsilon) = 1$$

(4) (b) Claim: $(W_n, n \in \mathbb{N})$ is a consistent sequence of estimators if

$$\forall \theta \in \Theta : \quad (i) \lim_{n\to\infty} Var_\theta W_n = 0$$

$$(ii) \lim_{n\to\infty} Bias_\theta W_n = 0$$

Proof: $P_\theta(|W_n - \theta| \geq \varepsilon) \overset{\text{Chebyshev}}{\leq} \frac{1}{\varepsilon^2} E_\theta(|W_n - \theta|^2)$ (1)

$= \frac{1}{\varepsilon^2}\left(Var_\theta(W_n - \theta) + (E_\theta(W_n - \theta))^2 \right)$ (1)

$= \frac{1}{\varepsilon^2}\left(Var_\theta(W_n) + (Bias_\theta(W_n))^2 \right)$ (1)

$\xrightarrow{n\to\infty} 0$ (1)

$\Rightarrow P_\theta(|W_n - \theta| < \varepsilon) \xrightarrow{n\to\infty} 1$ $\square$