Exam Statistical Inference Principles

LP 3, 2015/16

2016-03-14

$\|\Sigma 23\|$ Problem 1

- $(X_i, i=1,\dots,n)$ RS, $\mathcal{N}(\mu, \sigma^2)$-distributed
- $\overline{X} := \frac{1}{n}\sum_{i=1}^{n} X_i$ sample mean
- $S^2 := \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ sample variance

$\boxed{4}$ (a) Claim: The family of normal pdfs is an exponential family.

Proof:

Recall: A family of pdfs is called an exponential family if it can be expressed as

① $$f(x|\theta) = \underset{\geq 0}{h(x)}\, \underset{\geq 0}{c(\theta)}\, \exp\left(\sum_{i=1}^{k} w_i(\theta)\, t_i(x)\right)$$

Here $\theta = (\mu, \sigma^2)$ and

$$f(x|(\mu,\sigma^2)) \overset{\frac{1}{2}}{=} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x^2 - 2x\mu + \mu^2)\right)$$

$$\overset{\frac{1}{2}}{=} \underbrace{1}_{=h(x)\geq 0} \cdot \underbrace{(2\pi\sigma^2)^{-\frac{1}{2}}}_{=c(\mu,\sigma^2)\geq 0} \exp\left(\underbrace{\left(-\frac{1}{2\sigma^2}\right)}_{w_1(\mu,\sigma^2)} \underbrace{x^2}_{t_1(x)} + \underbrace{\frac{\mu}{\sigma^2}}_{w_2(\mu,\sigma^2)} \underbrace{x}_{t_2(x)} - \underbrace{\frac{\mu^2}{2\sigma^2}}_{w_3(\mu,\sigma^2)} \cdot \underbrace{1}_{t_3(x)} \right)$$

$4 \cdot \frac{1}{2} = 2$

$$= h(x)\, c(\mu,\sigma^2)\, \exp\left(\sum_{i=1}^{3} w_i(\mu,\sigma^2)\, t_i(x)\right)$$ □

$\boxed{6\frac{1}{2}}$ (b) Claim: $(\overline{X}, S^2)$ is minimal sufficient statistic for (μ, σ^2).

could be included in (ii) but then same computations there

Proof: (i) Show first sufficiency.

Recall the (Fisher-Neyman) Factorization Theorem:

$f(\cdot|\theta)$ joint pdf of RS $(X_i, i=1,\dots,n) =: X$

$T(X)$ sufficient statistic for θ

$\Leftrightarrow \exists g(\cdot|\theta), h: \forall x \in \underset{\text{sample space}}{\mathcal{X}}: \forall \theta \in \underset{\text{parameter space}}{\Theta}: f(x|\theta) = g(T(x)|\theta)\, h(x)$

Goal: Find g, h!

$$f((x_1,\dots,x_n)|(\mu,\sigma^2)) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right)$$

$$= (2\pi\sigma^2)^{-n/2} \exp\Big(-\frac{1}{2\sigma^2} \underbrace{\sum_{i=1}^{n}(x_i-\mu)^2}\Big)$$

$$= \sum_{i=1}^{n}(x_i-\bar{x}+\bar{x}-\mu)^2$$

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

$$= \sum_{i=1}^{n}(x_i-\bar{x})^2+2(\bar{x}-\mu)\underbrace{\Big(\sum_{i=1}^{n}x_i-n\bar{x}\Big)}_{=0} + n(\bar{x}-\mu)^2$$

$$= (2\pi\sigma^2)^{-n/2} \exp\Big(-\frac{1}{2\sigma^2}\Big(\sum_{i=1}^{n}(x_i-\bar{x})^2+n(\bar{x}-\mu)^2\Big)\Big)$$

$\frac{1}{2}$ $$\overset{\text{def}}{=} (2\pi\sigma^2)^{-n/2} \exp\Big(-\frac{1}{2\sigma^2}\big((n-1)s^2+n(\bar{x}-\mu)^2\big)\Big)$$

$$= g(\bar{x}, s^2)|(\mu,\sigma^2)$$

$\Rightarrow$ with $h(x)=1$:

$$f((x_1,\dots,x_n)|(\mu,\sigma^2)) = h(x)\, g((\bar{x},s^2)|(\mu,\sigma^2))$$

Factorization Theorem
$\Rightarrow$

$(\bar{X}, S^2)$ sufficient statistic for (μ,σ^2).

(ii) Show next minimality.

Recall the Theorem of Lehmann, Scheffé: 1

$f(\cdot|\theta)$ pdf of RS $(X_i, i=1,\dots,n)=X$

Assume: $\exists T: \forall x,y \in \mathcal{X}$

$$\frac{f(x|\theta)}{f(y|\theta)} = \text{const. as fction of } \theta \iff T(x)=T(y)$$

$\Rightarrow T(X)$ is minimal sufficient statistic.

Let $T(x) = (\bar{x}, s_x^2)$, then

$$\frac{f(x|(\mu,\sigma^2))}{f(y|(\mu,\sigma^2))} \overset{(i)}{=} \frac{(2\pi\sigma^2)^{-n/2}\exp\left(-\frac{1}{2\sigma^2}\left((n-1)s_x^2+n(\bar{x}-\mu)^2\right)\right)}{(2\pi\sigma^2)^{-n/2}\exp\left(-\frac{1}{2\sigma^2}\left((n-1)s_y^2+n(\bar{y}-\mu)^2\right)\right)}$$

$\frac{1}{2}$ + above

$$= \exp\left(-\frac{1}{2\sigma^2}\left((n-1)(s_x^2-s_y^2)+n\left(\bar{x}^2-2\bar{x}\mu+\mu^2-(\bar{y}^2-2\bar{y}\mu+\mu^2)\right)\right)\right)$$

$$= \exp\left(-\frac{1}{2\sigma^2}\left((n-1)(s_x^2-s_y^2)+n(\bar{x}^2-\bar{y}^2)-2n\mu(\bar{x}-\bar{y})\right)\right)$$

$\overset{!}{=}$ const. as fction of (μ, σ^2)

1

$$\Longleftrightarrow \quad (n-1)(s_x^2-s_y^2)+\underbrace{n(\bar{x}^2-\bar{y}^2)}_{=0} - \underbrace{2n\mu(\bar{x}-\bar{y})}_{\forall \mu \iff \bar{x}-\bar{y}=0} \overset{!}{=} 0$$

$\frac{1}{2}$ $\frac{1}{2}$

$$\Longleftrightarrow \quad \bar{x}-\bar{y} \overset{!}{=} 0 \quad \text{and} \quad (n-1)(s_x^2-s_y^2) \overset{!}{=} 0$$

$\frac{1}{2}$ $\frac{1}{2}$

$$\Longleftrightarrow \quad \bar{x}=\bar{y} \quad \text{and} \quad s_x^2 = s_y^2, \quad \text{i.e.,} \quad T(x)=T(y)$$

$\frac{1}{2}$ Lehmann, Scheffé

$$\Longrightarrow \quad (\bar{X}, S^2) \text{ minimal sufficient statistic} \qquad \square$$

$\boxed{6}$ (c) The likelihood function is given by

$$L((\mu,\sigma^2)\mid x) = \prod_{i=1}^{n} f(x_i \mid (\mu,\sigma^2))$$

1

$$\overset{(b)(i)}{=} (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\left((n-1)s^2+n(\bar{x}-\mu)^2\right)\right)$$

Compute

$$\frac{\partial}{\partial \sigma^2}\left(\log L((\mu,\sigma^2)\mid x)\right) = \frac{\partial}{\partial \sigma^2}\left(-\frac{n}{2}\log(2\pi\sigma^2)-\frac{1}{2\sigma^2}\left((n-1)s^2+n(\bar{x}-\mu)^2\right)\right)$$

$$\overset{\frac{1}{2}}{=} -\frac{n}{2}\frac{1}{2\pi\sigma^2}\cdot 2\pi + \frac{1}{2(\sigma^2)^2}\left((n-1)s^2+n(\bar{x}-\mu)^2\right)$$

$$\overset{\frac{1}{2}}{=} -\frac{n}{2}(\sigma^2)^{-1} + \left(\frac{(n-1)s^2}{2}+\frac{n}{2}\underbrace{(\bar{x}-\mu)^2}_{\substack{\| \\ 0 \text{ by assumption}}}\right)(\sigma^2)^{-2} \overset{!}{=} 0 \quad \text{to maximize}$$

$\frac{1}{2}$

$$\overset{\sigma^2 \neq 0}{\Longleftrightarrow} \quad -\frac{n}{2}\sigma^2 + \frac{(n-1)s^2}{2} = 0$$

$\frac{1}{2}$

$$\Longleftrightarrow \quad \sigma^2 = \frac{n-1}{n} S^2$$

1 Check that local maximum

$$\frac{\partial}{\partial(\sigma^2)^2}\left(\log L((\mu,\sigma^2)\mid x)\right) = \frac{n}{2}(\sigma^2)^{-2} - \left((n-1)s^2+\underbrace{(\bar{x}-\mu)^2}_{\substack{0 \text{ by assumption} \\ \|}}\right)(\sigma^2)^{-3}$$

$$\overset{\text{plug in}}{=} \frac{n}{2}\frac{n^2}{(n-1)^2}(S^2)^{-2} - \cancel{(n-1)}S^2\frac{n^3}{(n-1)^{\cancel{3}\,2}}(S^2)^{-3}$$

$$= -\frac{1}{2}\frac{n^3}{(n-1)^2}(S^2)^{-2} < 0 \Rightarrow \text{local maximum}$$

(1/2) $\lim_{\sigma^2 \to +\infty} \log \mathcal{L}((\mu,\sigma^2)|x) = -\infty$

(1/2) $\lim_{\sigma^2 \to 0} \log \mathcal{L}((\mu,\sigma^2)|x) = -\infty$ since $-\frac{1}{x}$ faster to $-\infty$ than $-\log x$ to $+\infty$

(1/2) $\Rightarrow \sigma^2 = \frac{n-1}{n} S^2$ unique maximum

$\Rightarrow$ since log likelihood has the same maximum as the likelihood,

(1/2) $\hat{\sigma}^2 := \frac{n-1}{n} S^2$ is the maximum likelihood estimator of σ^2.

$\boxed{4\frac{1}{2}}$

(d) The <u>mean squared error</u> of an estimator W of a parameter θ is the function of θ defined by $\mathbb{E}_\theta((W-\theta)^2)$. (1)

$$MSE(S^2) = \mathbb{E}_{\mu,\sigma^2}((S^2-\sigma^2)^2) \overset{\text{since unbiased}}{=} Var_{\mu,\sigma^2}(S^2) = \frac{2\sigma^4}{n-1}$$

(1/2) (1/2)

$$MSE(\hat{\sigma}^2) = \mathbb{E}_{\mu,\sigma^2}((\hat{\sigma}^2-\sigma^2)^2) \overset{(1/2)}{=} Var_{\mu,\sigma^2}(\hat{\sigma}^2) + (Bias_{\mu,\sigma^2}(\hat{\sigma}^2))^2$$

$$= Var_{\mu,\sigma^2}\left(\frac{n-1}{n}S^2\right) + \left(Bias_{\mu,\sigma^2}\left(\frac{n-1}{n}S^2\right)\right)^2$$

(1/2) (1/2)

$$= \left(\frac{n-1}{n}\right)^2 \frac{2\sigma^4}{n-1} + \left(\mathbb{E}_{\mu,\sigma^2}\left(\frac{n-1}{n}S^2\right) - \sigma^2\right)^2$$

(1/2)

$$\overset{S^2 \text{ unbiased}}{=} \frac{2(n-1)}{n^2}\sigma^4 + \left(\frac{n-1}{n}\sigma^2 - \sigma^2\right)^2$$

$$= \frac{2(n-1)}{n^2}\sigma^4 + \frac{1}{n^2}\sigma^4$$

(1/2)

$$= \frac{2n-1}{n^2}\sigma^4$$

$\boxed{2}$

(e) S^2 "better" than $\hat{\sigma}^2$ because unbiased and therefore on average correct (1)

$\hat{\sigma}^2$ "better" than S^2 because $MSE(\hat{\sigma}^2) < MSE(S^2)$ (1)

$\boxed{\Sigma 13}$ Problem 2

$\boxed{6}$ (a) Claim:

Let f be a unimodal pdf. If the interval $[a,b]$ satisfies

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

(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 amoung all intervals that satisfy.

Proof: Assume that we have shown the claim for all intervals $[a',b']$ such that $a' \le a$.

1. Assume that there exists an interval $[a',b']$ that satisfies (i) and that is shorter than $[a,b]$, where $a' > a$

case $b' \le b$

$$\int_{a'}^{b'} f(x)\,dx \overset{\frac{1}{2}}{=} \underbrace{\int_a^b f(x)\,dx}_{=1-\alpha} - \underbrace{\int_a^{a'} f(x)\,dx}_{\substack{\text{since } f(a)>0 \\ \text{and } f(x)\ge f(a) \\ \forall x\in[a,a'] \\ >0}} - \underbrace{\int_{b'}^{b} f(x)\,dx}_{\ge 0}$$

$\frac{1}{2}$ $\frac{1}{2}$

$$< 1-\alpha$$

↯

case $b' > b$, $a' \le b$

$$\int_{a'}^{b'} f(x)\,dx \overset{\frac{1}{2}}{=} \int_a^b f(x)dx - \int_a^{a'} f(x)dx + \int_b^{b'} f(x)\,dx$$

$$\overset{f \text{ unimodal}}{\le} 1-\alpha - f(a)(a'-a) + f(b)(b'-b) \quad \frac{1}{2}$$

$$\overset{f(a)=f(b)}{=} 1-\alpha + \underbrace{f(a)}_{\substack{>0 \\ (ii)}}\underbrace{\left((b'-a') - (b-a)\right)}_{\substack{\text{shorter} \\ <0}} \quad \frac{1}{2}$$

$\frac{1}{2}$

$$< 1-\alpha$$

↯

case $b < a' < b'$

$$\int_{a'}^{b'} f(x)\,dx \overset{\text{unimodal}}{\underset{\left(\frac{1}{2}\right)}{\leq}} f(a')(b'-a')$$

$$\underset{\substack{b'-a'<b-a \\ f(b)>0}}{\overset{\left(\frac{1}{2}\right)}{\leq}} f(b)(b'-a')$$

$$< f(b)(b-a)$$

$$\overset{\text{unimodal}}{\underset{\left(\frac{1}{2}\right)}{\leq}} \int_a^b f(x)\,dx = 1-\alpha \quad ⚡ \quad \square$$

$\boxed{2\frac{1}{2}}$

(b) Let $f(x) = \mathbb{1}_{[0,1]}(x)$, then f is the pdf of the uniform distribution on $[0,1]$ and for all $x^* \in [0,1]$ it is nondecreasing on $(-\infty, x^*]$ and nonincreasing on $[x^*, +\infty)$. Then for fixed $\alpha > 0$ and all $a, b \in [0,1]$ s.t. $b-a = 1-\alpha$: (1)

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

$\left(\frac{1}{2}\right)$

(but also (ii) and (iii) are fullfilled). There exists infinitely many a, b, eg. $a = 0$, $b = 1-\alpha$ or $a = \frac{\alpha}{2}$, $b = 1 - \frac{\alpha}{2}$. More precisely

$$\forall a \in [0,\alpha] : \exists b \in [0,1] : \int_a^b f(x)\,dx = 1-\alpha,$$

actually $b = 1-\alpha+a$.

$\boxed{4\frac{1}{2}}$

(c) • $(X_i, i=1,\dots,n)$ RS, $\mathcal{N}(\mu, \sigma^2)$ distributed

• $\bar{X}$ sample mean, which is $\mathcal{N}(\mu, \frac{\sigma^2}{n})$ distributed

$$\Rightarrow \quad Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1) \text{ distributed} \quad (1)$$

Let $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ be the corresponding density, then

• $x^* = 0$ is a mode $\left(\frac{1}{2}\right)$

• f is symmetric around 0 $\left(\frac{1}{2}\right)$

$\Rightarrow f(a) = f(b) > 0$ for $a = -b$ $\left(\frac{1}{2}\right)$

and $-b \leq x^* = 0 \leq b$ $\left(\frac{1}{2}\right)$

Choose b such that $\int_{-b}^{b} f(x)\,dx = 1-\alpha = P(|Z| \leq b)$ (exists!) $\left(\frac{1}{2}\right)$

$\overset{(a)\ \left(\frac{1}{2}\right)}{\Rightarrow}$ $[-b, b]$ is a shortest $1-\alpha$ confidence interval for 0 $\left(\frac{1}{2}\right)$

$\overset{\text{transform}}{\Rightarrow}$ $\{\mu \in \mathbb{R}, -b \leq z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \leq b\} = \{\mu \in \mathbb{R}, \bar{x} - b\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + b\frac{\sigma}{\sqrt{n}}\}$ shortest $1-\alpha$ confidence interval

Σ16 | Problem 3

$X_i \sim \mathcal{N}(\theta, \sigma^2)$, $\theta \sim \mathcal{N}(\mu, \tau^2)$, $\pi(\cdot|\bar{x}) \sim \mathcal{N}(\mathbb{E}(\theta|\bar{x}), \mathrm{Var}(\theta|\bar{x}))$

3 (a) The Baysian approach to statistics assumes that the parameter θ, which should be estimated by a random sample $X = (X_i, i=1,\dots,n)$, is a random variable ½ with so called prior distribution ½ $\pi(\theta)$. After the observation x it is updated to the posterior distribution ½ $\pi(\theta|x)$ according to Bayes' Rule:

$$\pi(\theta|x) := \frac{f(x|\theta)\pi(\theta)}{m(x)}, \quad \tfrac{1}{2}$$

where ½ $f(\cdot|\theta)$ is the pdf/pmf of X and $m(x) = \int_\Theta$ ½ $f(x|\theta)\pi(\theta)\,d\theta$ (in the cont case) is the marginal distribution.

2 (b) Let $\mathcal{F}$ denote a/the class of pdfs or pmfs $f(\cdot|\theta)$. A class Π of prior distributions is a conjugate family for $\mathcal{F}$ if the posterior distribution is in the class Π for all $f \in \mathcal{F}$, all priors in Π and all $x \in \mathcal{X}$. 1

1 An example is that the normal distribution is conjugate to itself which we computed in the exercises.

2 (c) A possible choice of a point estimator is the mean of the 1 posterior, i.e.,

1 $$\mathbb{E}(\theta|\bar{x}) = \int_\Theta \theta\, \pi(\theta|\bar{x})\,d\theta \overset{\text{exercise/given}}{=} \frac{\tau^2}{\tau^2+\sigma^2/n}\bar{x} + \frac{\sigma^2/n}{\tau^2+\sigma^2/n}\mu$$

4 (d) For testing $H_0: \theta \le \theta_0$ vs. $H_1: \theta > \theta_0$ and observation x, we have posterior probabilities

½ $P(\theta \le \theta_0|\bar{x}) = P(H_0 \text{ is true}|\bar{x})$

½ $P(\theta > \theta_0|\bar{x}) = P(H_1 \text{ is true}|\bar{x})$

With possible decision (but others are also fine)

½ accept H_0 if $P(\theta \le \theta_0|\bar{x}) \ge P(\theta > \theta_0|\bar{x})$,

½ i.e., $P(\theta \le \theta_0|\bar{x}) \ge \frac{1}{2}$

$\frac{1}{2}$

Since normal distribution symmetric, i.e., $P(\Theta \leq E(\Theta)) = \frac{1}{2}$:

$$P(\theta \leq \theta_0 \mid \bar{x}) \geq \frac{1}{2}$$

$\frac{1}{2}$ $\iff$ $\mathbb{E}(\theta \mid \bar{x}) \leq \theta_0$

$$\mathbb{E}(\theta \mid \bar{x}) = \frac{\tau^2 \bar{x} + \frac{\sigma^2}{n}\mu}{\tau^2 + \frac{\sigma^2}{n}} = \frac{n\tau^2\bar{x} + \sigma^2\mu}{n\tau^2 + \sigma^2}$$

$\frac{1}{2}$ $\iff$ $\bar{x} \leq \frac{(n\tau^2 + \sigma^2)\theta_0 - \sigma^2\mu}{n\tau^2} = \theta_0 + \frac{\sigma^2(\theta_0 - \mu)}{n\tau^2}$

The acceptance region is given by

$\frac{1}{2}$ $\left\{ x \in \mathcal{X},\ \bar{x} \leq \theta_0 + \frac{\sigma^2(\theta_0 - \mu)}{n\tau^2} \right\}$.

$\boxed{5}$

(e) If $\pi(\cdot \mid x)$ is the posterior distribution, then for any $A \subset \Theta$,

$P(\theta \in A \mid X = x) = \int_A \pi(\theta \mid x)\, d\theta$ is the credible probability and

1 A is a credible set.

Observe that $\frac{\theta - \mathbb{E}(\theta \mid \bar{x})}{\sqrt{Var(\theta \mid \bar{x})}} \sim \mathcal{N}(0,1)$, then take $b > 0$ such $\frac{1}{2}$

that $\frac{1}{2}$ $\int_{-b}^{b} \frac{1}{\sqrt{2\pi}} e^{-\frac{\theta^2}{2}}\, d\theta = 1 - \alpha$ (see Problem 2 (c))

$\Rightarrow A \overset{\frac{1}{2}}{=} \left\{ \theta \in \Theta,\ \left| \frac{\theta - \mathbb{E}(\theta \mid \bar{x})}{\sqrt{Var(\theta \mid \bar{x})}} \right| \leq b \right\}$

$= \left\{ \theta \in \Theta,\ -b \leq \frac{\theta - \mathbb{E}(\theta \mid \bar{x})}{\sqrt{Var(\theta \mid \bar{x})}} \leq b \right\}$

$\overset{\frac{1}{2}}{=} \left\{ \theta \in \Theta,\ \mathbb{E}(\theta \mid \bar{x}) - b\sqrt{Var(\theta \mid \bar{x})} \leq \theta \leq \mathbb{E}(\theta \mid \bar{x}) + b\sqrt{Var(\theta \mid \bar{x})} \right\}$

where the parameters as above, is a $1-\alpha$ credible set for θ,

1 i.e., the experimenter is $(1-\alpha)\%$ sure that θ is in the interval.

By construction with the result from Problem 2(a), we have

1 constructed the shortest credible interval, i.e., the highest posterior density region. (see also Corollary 9.3.10 in the lecture).

Σ8 Problem 4

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

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

i.e., $(W_n, n\in\mathbb{N})$ converges in probability to θ.

5 (b) Claim:

- $(W_n, n\in\mathbb{N})$ consistent sequence of estimators of θ
- $(a_n, n\in\mathbb{N})$, $(b_n, n\in\mathbb{N})$ sequences with
  (i) $\lim_{n\to\infty} a_n = 1$
  (ii) $\lim_{n\to\infty} b_n = 0$

$\Rightarrow$ $(a_nW_n+b_n, n\in\mathbb{N})$ consistent sequence of estimators of θ

1/2 (or last conclusion)

Proof: Show that $(a_nW_n+b_n, n\in\mathbb{N})$ converges in probability to θ

1 $\varepsilon'\in(0,\varepsilon) : \exists n_0\in\mathbb{N} : \forall n>n_0 : |a_n-1||\theta|+|b_n|<\varepsilon'$ and $|a_n|<1+\varepsilon'$,

which exists due to (i) and (ii)

Then:

$$\{\omega\in\Omega,\ |a_nW_n(\omega)+b_n-\theta|<\varepsilon\}$$

$$\overset{1/2}{=}\{\omega\in\Omega,\ |a_n(W_n(\omega)-\theta)+(a_n-1)\theta+b_n|<\varepsilon\}$$

$$\overset{1/2}{\supseteq}\{\omega\in\Omega,\ |a_n||W_n(\omega)-\theta|+|a_n-1||\theta|+|b_n|<\varepsilon\}$$

$$\overset{1/2}{\underset{n>n_0}{\supseteq}}\{\omega\in\Omega,\ |W_n(\omega)-\theta|<\underbrace{\frac{\varepsilon-\varepsilon'}{1+\varepsilon'}}_{=:\varepsilon''>0}\}$$

$$\overset{1/2}{=}\{\omega\in\Omega,\ |W_n(\omega)-\theta|<\varepsilon''\}$$

$$\Rightarrow \lim_{n\to+\infty} P_\theta(|a_nW_n+b_n|<\varepsilon) \overset{1/2}{\geq} \lim_{n\to+\infty} P_\theta(|W_n-\theta|<\varepsilon'') \overset{1/2}{=} 1$$

$$\overset{1/2}{\underset{\text{prob}\in[0,1]}{\Rightarrow}} \lim_{n\to+\infty} P_\theta(|a_nW_n+b_n|<\varepsilon)=1,$$

i.e., by definition of a consistent sequence the claim □

2

(c) Consistency cares just about the limit but all finite estimators can be (arbitrarily) bad, which can be achieved by scaling of the first n_0 a_n and b_n. 1

=> Just for (very) large sample sizes an indication. 1