Annika Lang (1)

Statistical Inference Principles

Exam: March 10, 2014, 14:00 – 18:00

Solutions

$\boxed{\Sigma 22}$ **Problem 1**

$(X_i, i=1,\dots,n)$ RS, $\mathcal{N}(\theta, \sigma^2)$ distributed, σ^2 known

$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ sample mean

($\Sigma 4$) (a) Claim: $\bar{X}$ is a minimal sufficient statistic for μ, σ^2 known.

Proof

not needed

sufficiency: We know (Thm. 6.2.2) $\bar{X}$ sufficient iff

$$\forall x \in \mathcal{X}: \frac{p(x|\theta)}{q(\bar{x}|\theta)} \text{ constant as function of } \theta$$

where p joint pdf of X and q pdf of $\bar{X}$

Compute

$$p(x|\theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} = (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\theta)^2}$$

$$= (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left((x_i-\bar{x})^2 - 2(x_i-\bar{x})(\bar{x}-\mu) + (\bar{x}-\theta)^2\right)}$$

$$= (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n}(x_i-\bar{x})^2 - 0 + n\cdot(\bar{x}-\theta)^2\right)}$$

and since $\bar{X} \sim \mathcal{N}(\theta, \frac{\sigma^2}{n})$

$$q(\bar{x}|\theta) = \frac{1}{\sqrt{2\pi\sigma^2/n}} e^{-\frac{(\bar{x}-\theta)^2}{2\sigma^2/n}}$$

$$\Rightarrow \frac{p(x|\theta)}{q(\bar{x}|\theta)} = \frac{(2\pi\frac{\sigma^2}{n})^{1/2}}{(2\pi\sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\bar{x})^2}$$ constant for fixed x and all $\theta \in \mathbb{R}$

$\Rightarrow \bar{X}$ is sufficient statistic.

minimality (+ sufficiency): (Thm. 6.2.13, Lehmann/Scheffé)

(1) $f(\cdot|\theta)$ pdf and $T: \mathcal{X} \to \mathbb{R}$

Assume $\forall x, y \in \mathcal{X}: \frac{f(x|\theta)}{f(y|\theta)}$ const. as function of $\theta \iff T(x) = T(y)$

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

(1) It holds that

$$\frac{f(x|\theta)}{f(y|\theta)} = \frac{(2\pi\sigma^2)^{-n/2}}{(2\pi\sigma^2)^{-n/2}} \frac{e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\theta)^2}}{e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\theta)^2}} = e^{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n}(x_i^2-y_i^2) + 2\theta\left(\sum_{i=1}^{n}x_i-y_i\right)\right)}$$

$$= e^{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n}(x_i^2-y_i^2)\right)} e^{\frac{\theta}{\sigma^2}n\cdot(\bar{x}-\bar{y})}$$

and $e^{-\frac{\theta}{\sigma^2}n(\bar{x}-\bar{y})}$ is constant for all θ $\Leftrightarrow$ $\bar{x}-\bar{y}=0 \Leftrightarrow \bar{x}=\bar{y}$ (1)

Lehmann/Scheffé
$\Rightarrow$ $\bar{X}$ minimal sufficient statistic for θ. $\square$ (1)

(b) Claim: The family of normal pdfs with known variance σ^2 is an exponential family (Σ2)

Proof: Set $h(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \geq 0$, $c(\theta) = e^{-\frac{\theta^2}{2\sigma^2}}$, $w(\theta)=\frac{\theta}{\sigma^2}$, $t(x)=x$ (1)

then

$$f(x|\theta) = h(x)\,c(\theta)\exp(w(\theta)t(x))$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x^2+\theta^2)} e^{\frac{\theta x}{\sigma^2}}$$ (1)

$$= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x^2-2x\theta+\theta^2)} = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\theta)^2}{2\sigma^2}}$$

the pdf of a $\mathcal{N}(\theta,\sigma^2)$ RV

$\Rightarrow \mathcal{N}(\theta,\sigma^2)$ is exponential family, since splitted according to definition. $\square$

(c) Claim: $\bar{X}$ is a complete statistic for θ (Σ4)

Proof: From Thm. 6.2.25 (complete statistics in the exponential families) (1)

we know that if the RS has a distribution from an exponential family,

then $T(X) = \sum_{i=1}^{n} X_i$ is complete if $\{w(\theta), \theta\in\Theta\}$ open set in $\mathbb{R}$

We have from (b) that $w(\theta)=\frac{\theta}{\sigma^2}$, $\theta\in\mathbb{R}$ $\Rightarrow$ $\{w(\theta), \theta\in\Theta\}=\mathbb{R}$ open in $\mathbb{R}$ (1)

So $T(X) = \sum_{i=1}^{n} X_i = n\bar{X}$ is a complete statistic (1)

Since scaling does not affect completeness ($\mathbb{E}_\theta(g(\bar{X})) = \mathbb{E}_\theta(g'(\sum_{i=1}^{n} X_i))$ (1)

with $g'(x) := g(\frac{1}{n}x)$), $\bar{X}$ is a complete statistic. $\square$

(d) The method of moments states: (Σ5)

$$m_1 = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \mu_1 = \mathbb{E}(X_1) = \theta$$ (1)

$$m_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 \qquad \mu_2 = \mathbb{E}(X_1^2) = \theta^2+\sigma^2$$

$\Rightarrow$ $\hat{\theta} = m_1 = \bar{X}$ (1)

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - (\bar{X})^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2 - \frac{1}{n}\Big(\underbrace{\sum_{i=1}^{n}\bar{X}^2}_{n\bar{X}^2} - 2\bar{X}\underbrace{\sum_{i=1}^{n}X_i}_{n\bar{X}^2}\Big) - \bar{X}^2$$ (1)

$$= \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

$\mathbb{E}(\hat{\theta}) = \mathbb{E}(\bar{X}) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}(X_i) = \theta$ $\Rightarrow$ $\hat{\theta}$ unbiased estimator (1)

$\mathbb{E}(\hat{\sigma}^2) = \frac{n}{n-1}\mathbb{E}(S^2) = \frac{n}{n-1}\sigma^2$ $\Rightarrow$ $\hat{\sigma}^2$ not unbiased estimator since $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$ is (1)

(e) The likelihood function is $\Sigma 4$

$$\mathcal{L}(\theta|x) = \prod_{i=1}^{n} f(x|\theta) = (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\theta)^2}$$

Then the maximum likelihood estimator is (1)

$$MLE = \hat{\theta} = \underset{\theta\in\Theta}{\operatorname{argmax}}\ \mathcal{L}(\theta|X)$$

Observe that $\sum_{i=1}^{n}(x_i-\theta)^2 \geq \sum_{i=1}^{n}(x_i-\bar{x})^2 \quad \forall\theta\in\mathbb{R}$

and $\sum_{i=1}^{n}(x_i-\theta)^2 = \sum_{i=1}^{n}(x_i-\bar{x})^2 \iff \theta=\bar{x}$ (2) comp.

$$\Rightarrow e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\theta)^2} \leq e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

$$\Rightarrow MLE = \hat{\theta} = \mathcal{L}(\bar{X}|X)$$ (1)

(f) The mean squared error of an estimator W of a parameter θ is the $\Sigma 3$ (1) function of θ defined by $\mathbb{E}_\theta((W-\theta)^2)$.

For $W=\bar{X}$ it holds that

$$\mathbb{E}_\theta((\bar{X}-\theta)^2) \overset{\bar{X}\text{ unbiased}}{=} \mathbb{E}_\theta((\bar{X}-\mathbb{E}_\theta(\bar{X}))^2) = Var_\theta(\bar{X}) = \sigma^2/n$$ (2)

$\boxed{\Sigma 9}$ Problem 3

$(X_i, i=1,\ldots,n)$ RS, $\mathcal{N}(\theta,\sigma^2)$ distributed, σ^2 known

$\Sigma 4$ (a) (1) A pivot or pivotal quantity is a RV $Q(X,\theta)$ that has the same distribution for all $\theta\in\Theta$.

(1) We know that $\bar{X}\sim\mathcal{N}(\theta,\sigma^2/n)$, therefore $\bar{X}-\theta\sim\mathcal{N}(0,\sigma^2/n)$

(1) $\Rightarrow \bar{X}-\theta$ has pdf $\frac{1}{\sqrt{2\pi\sigma^2/n}}e^{-\frac{x^2}{2\sigma^2/n}}$ which does not depend on θ

(1) $\Rightarrow \bar{X}-\theta$ is pivot for θ

$\Sigma 5$ (b) Let $a<b$ s.t.

(1) + (1) exists. $P_\theta(a\leq\bar{X}-\theta\leq b) = 1-\alpha$, where $\bar{X}-\theta\sim\mathcal{N}(0,\sigma^2/n)$, so that a,b exist

(1) $\Rightarrow$ Set $C(x)=\{\theta\in\mathbb{R}, \bar{x}-b\leq\theta\leq\bar{x}-a\}$, then $C(X)$ is a $1-\alpha$ confidence interval

(1) cite Thm. 9.3.2

(1) $\Rightarrow$ Choosing $a=-b$ leads to the shortest $1-\alpha$ confidence interval

Problem 2

$\boxed{\Sigma 16}$

$(X_i, i=1,\dots,n)$ RS, $\mathcal{N}(\theta, \sigma^2)$ distributed, σ^2 known

Test $H_0: \theta \leq \theta_0$ vs. $H_1: \theta > \theta_0$

(a) The likelihood ratio statistic λ for testing H_0 vs. H_1 is given by $\Sigma 2$

$$\lambda(x) := \frac{\sup_{\theta \leq \theta_0} L(\theta | x)}{\sup_{\theta \in \mathbb{R}} L(\theta | x)}, \quad x \in \mathcal{X}$$ 1

A likelihood ratio test is any test that has a rejection region of the form $\{x \in \mathcal{X}, \lambda(x) \leq c\}$ for some $c \in [0,1]$. 1

(b) Observe that $\sup_{\theta \leq \theta_0} L(\theta | x) = \sup_{\theta \leq \theta_0} (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i - \theta)^2}$ $\Sigma 4$

$$= (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i - \min\{\theta_0, \bar{x}\})}$$

$\bar{x}$ leads to maximum, but if $\bar{x} > \theta_0$, choose "nearest" to $\bar{x}$. 1

$$\Rightarrow \lambda(x) = \frac{(2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i - \min\{\theta_0, \bar{x}\})^2}}{(2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i - \bar{x})^2}} = \begin{cases} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n ((x_i - \theta_0)^2 - (x_i - \bar{x})^2)} & \text{if } \bar{x} > \theta_0 \\ 1 & \text{if } \bar{x} \leq \theta_0 \end{cases}$$

$$\overset{\text{(calc. 1(a) suff.)}}{=} \begin{cases} e^{-\frac{1}{2\sigma^2} n(\bar{x} - \theta_0)^2} & \text{if } \bar{x} > \theta_0 \\ 1 & \text{if } \bar{x} \leq \theta_0 \end{cases}$$ 1

$\Rightarrow$ rejection region

$$\mathcal{R}(x) = \{x \in \mathcal{X}, \lambda(x) \leq c\}$$

else $\mathcal{R}(x) = \mathcal{X}$ $\leftarrow$

$$\overset{c<1}{=} \{x \in \mathcal{X}, \bar{x} > \theta_0, e^{-\frac{1}{2\sigma^2} n(\bar{x} - \theta_0)^2} \leq c\}$$ 2

$$= \{x \in \mathcal{X}, \bar{x} > \theta_0, (\bar{x} - \theta_0)^2 \geq -\frac{2\sigma^2}{n} \log c\}$$

$$= \{x \in \mathcal{X}, \bar{x} > \theta_0, \bar{x} - \theta_0 \geq \sqrt{-\frac{2\sigma^2}{n} \log c}\}$$

(c) A family of pdfs or pmfs $\{g(\cdot | \theta), \theta \in \Theta\}$ for a univariate RV with real-valued parameter θ has a monotone likelihood ratio if $\Sigma 3$ 1

$$\forall \theta_2 > \theta_1: \frac{g(\cdot | \theta_2)}{g(\cdot | \theta_1)} \text{ is a monotone function on } \{t \in \mathcal{T}, g(t | \theta_1) > 0 \text{ or } g(t | \theta_2) > 0\},$$

where $\frac{c}{0} := +\infty$ for $c > 0$, ($\mathcal{T}$ is the "sample space" of T).

Let $T \sim \mathcal{N}(\theta, \sigma^2)$, σ^2 known, and $\theta_2 > \theta_1$, then for $t \in \mathbb{R}$:

$$\frac{g(t | \theta_2)}{g(t | \theta_1)} = \frac{(2\pi\sigma^2)^{-1/2} e^{-\frac{1}{2\sigma^2}(t - \theta_2)^2}}{(2\pi\sigma^2)^{-1/2} e^{-\frac{1}{2\sigma^2}(t - \theta_1)^2}} = e^{-\frac{1}{2\sigma^2}((t - \theta_2)^2 - (t - \theta_1)^2)} = e^{-\frac{1}{2\sigma^2}(2t(\theta_1 - \theta_2) + \theta_2^2 - \theta_1^2)}$$

$$= e^{-\frac{\theta_2^2 - \theta_1^2}{2\sigma^2}} e^{\overbrace{\frac{\theta_2 - \theta_1}{\sigma^2}}^{>0} t}, \text{ increasing function in } t$$ 2

CHALMERSSKA
HUSET

$\Sigma 5$

(d) For $\alpha \in [0,1]$, a <u>level α test</u> is a test such that $P_\theta(X \in R) \le \alpha$, where R denotes the rejection region of the test. (1)

Within the class of level α tests, a test is a <u>uniformly most powerful test</u> if $P_\theta(X \in R) \ge P_\theta(X \in R') \ \forall \theta > \theta_0 \ (\theta \in \Theta_0^c)$, where R' is the rejection region of any level α test. (1)

By the Karlin–Rubin theorem, the test "reject $H_0 \iff T(X) > t_0$" with $P_{\theta_0}(T(X) > t_0)$ is a UMP level α test if $T(X)$ is a sufficient statistic for θ (1)

We know (Pr. 1(a)) that $\bar{X}$ is a sufficient statistic for θ (1)

$\overset{\text{Karlin–Rubin}}{\Longrightarrow}$ Since $\bar{X} \sim N(\theta, \sigma^2/n)$, t_0 with $P_{\theta_0}(\bar{X} > t_0) = \alpha$ can be computed. Then

reject $H_0 \iff \bar{X} > t_0$ (1)

is a UMP level α test.

(e) We know (Thm. 8.3.27) that $p(X)$ is a valid p-value if $\Sigma 2$

$$p(x) := \sup_{\theta \le \theta_0} P_\theta(W(X) \ge W(x)), \quad x \in \mathcal{X},$$

where $W(X)$ is a test statistic s.t. large values of $W(X)$ give evidence that H_1 is true. (1)

Set $W(X) = \bar{X}$, then large values of $\bar{X}$ ($> \theta_0$) give evidence that H_1 is true. Therefore a valid p-value can be defined by

$$p(x) = \sup_{\theta \le \theta_0} P_\theta(\bar{X} \ge \bar{x}) = \int_{\bar{x}}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2/n}} e^{-\frac{1}{2\sigma^2/n} \min_{\theta \le \theta_0} \{(y-\theta)^2\}} dy$$ (1)

$$= \int_{\bar{x}}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2/n}} e^{-\frac{1}{2\sigma^2/n} \begin{cases} (y-\theta_0)^2 & \text{if } \bar{x} > \theta_0 \\ (y-\bar{x})^2 & \text{else} \end{cases}} dy$$

Problem 4

$\Sigma 13$

$\Sigma 5$

(a) An estimator W is a <u>best unbiased estimator</u> of $\tau(\theta)$ if $E_\theta(W) = \tau(\theta)$ for all $\theta \in \Theta$ and, for any other estimator W^* with $E_\theta(W^*) = \tau(\theta)$, $Var_\theta(W) \le Var_\theta(W^*)$ for all $\theta \in \Theta$. (1)

<u>Claim</u>: W best unbiased estimator of $E_\theta(W)$

$\Rightarrow \ \forall$ unbiased estimators U of 0: $Cov_\theta(W, U) = 0$

Proof:

W best unbiased estimator

$\Rightarrow$ for U unbiased estimator of 0 : $\mathbb{E}_\theta(W + aU) = \mathbb{E}_\theta(W)$, i.e. unbiased

$$Var_\theta(W + aU) = Var_\theta(W) + a^2 Var_\theta(U) + 2a\, Cov_\theta(W, U) \overset{!}{\underset{\text{since best}}{\geq}} Var_\theta(W)$$ 1

$$\Rightarrow \forall a \in \mathbb{R} : a^2 Var_\theta(U) + 2a\, Cov_\theta(W, U) \overset{!}{\geq} 0$$

But: Assume that $Cov_{\theta_0}(W, U) \neq 0$ for some θ_0 1

then
$$a^2 Var_{\theta_0}(U) + 2a\, Cov_{\theta_0}(W, U) < 0$$
$$\Leftrightarrow a^2 + 2a \frac{Cov_{\theta_0}(W, U)}{Var_{\theta_0}(U)} < 0$$
$$\Leftrightarrow a \left(a + 2 \frac{Cov_{\theta_0}(W, U)}{Var(U)}\right) < 0$$

for $\frac{Cov_\theta(W, U)}{Var(U)} > 0 \Rightarrow a \in \left(-2 \frac{Cov_{\theta_0}(W, U)}{Var_{\theta_0}(U)}, 0\right)$ 1

for $\frac{Cov_{\theta_0}(W, U)}{Var(U)} < 0 \Rightarrow a \in \left(0, -2 \frac{Cov_{\theta_0}(W, U)}{Var_{\theta_0}(U)}\right)$ 1

i.e.,
$$Var_{\theta_0}(W + aU) < Var_{\theta_0}(W)$$ ↯ W best unbiased estimator.

(b) $\forall \theta_0 \in \Theta$: $A(\theta_0)$ acceptance region of level α test of Σ3

$$H_0 : \theta = \theta_0 \text{ vs. } H_1 : \theta \neq \theta_0$$

Set
$$C(x) := \{\theta_0 \in \Theta,\ x \in A(\theta_0)\}, \quad x \in \mathcal{X}$$

Claim: $C(X)$ is a $1-\alpha$ confidence set

Proof: We have to show that $P_\theta(\theta \in C(X)) \geq 1 - \alpha$ 1

For $R(\theta_0)$ the corresponding rejection regions, it holds that

$\alpha \leq P_{\theta_0}(X \in R(\theta_0)) = 1 - P_{\theta_0}(X \in A(\theta_0))$ 1

$\Leftrightarrow 1 - \alpha \geq P_{\theta_0}(X \in A(\theta_0)) \overset{\text{inversion}}{=} P_{\theta_0}(\theta_0 \in C(X))$ 1

□

CHALMERSSKA
HUSET

Σ5

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

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

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

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