

Optimal procedures (decision theory cont.)

Lecture 18

Equivariance Equivalent ~~estimated~~ decisions, invariant decision.

Unbiasedness To define this..

Assume L loss function satisfies if d is correct decision for two different values θ_1, θ_2 then assume $L(\theta_1, d) = L(\theta_2, d)$

Define a decision rule is λ -unbiased if

$$(1) \quad E_{\theta}(L(\theta, \delta(\bar{x}))) \leq E_{\theta'}(L(\theta', \delta(\bar{x})))$$

for $\theta \neq \theta'$.

Ex For confidence interval construction the decision rule is

$$\delta(\bar{x}) = (l(\bar{x}), u(\bar{x}))$$

Then, if the loss function is

$$L(\theta, d) = 1 \{ \theta \notin d \},$$

①

then δ decision being λ -unbiased means

$$\begin{aligned} & E_{\theta} (1_{\{\theta \notin \delta(\bar{x})\}}) \leq E_{\theta} (1_{\{\theta' \notin \delta(\bar{x})\}}) \\ (\Leftarrow) \quad & P_{\theta} (\theta \notin \delta(\bar{x})) \leq P_{\theta} (\theta' \notin \delta(\bar{x})) \\ (\Leftarrow) \quad & P_{\theta} (\theta \in \delta(\bar{x})) \geq P_{\theta} (\theta' \in \delta(\bar{x})) \end{aligned}$$

~~#~~

Note θ is the correct parameter, meaning that $\bar{x} \sim F_{\theta}$, θ' is any (other) parameter.

#

Ex. Hypothesis testing $\mathcal{D} = \{d_0, d_1\}$
and $\mathcal{N} = \mathcal{R}_0 \cup \mathcal{R}_1$,

d_0 decision that $\theta \in \mathcal{R}_0$
 d_1 - if - $\theta \in \mathcal{R}_1$

introduce a loss function (as before)
that is 0 if correct decision made

(2)

and a_0 when wrong decision made for $\theta \in \mathcal{R}_0$, and a_1 when wrong decision is made for $\theta \in \mathcal{R}_1$, i.e.

$$L(\theta, d) = a_1 \cdot 1\{d = d_0\} \cdot 1\{\theta \in \mathcal{R}_1\} \\ + a_0 \cdot 1\{d = d_1\} \cdot 1\{\theta \in \mathcal{R}_0\}$$

so that an L-unbiased decision rule satisfies

$$\mathbb{E}_{\theta} (L(\theta, \delta(\xi))) \leq \mathbb{E}_{\theta'} (L(\theta', \delta(\xi)))$$

$$a_1 P_{\theta} (\delta(\xi) = d_0) \cdot 1\{\theta \in \mathcal{R}_1\} + a_0 P_{\theta} (\delta(\xi) = d_1) \cdot 1\{\theta \in \mathcal{R}_0\}$$

$$\leq a_1 P_{\theta'} (\delta(\xi) = d_0) \cdot 1\{\theta' \in \mathcal{R}_1\} + a_0 P_{\theta'} (\delta(\xi) = d_1) \cdot 1\{\theta' \in \mathcal{R}_0\}$$

This becomes

$$a_1 P_{\theta} (\delta(\xi) = d_0) \geq a_0 P_{\theta} (\delta(\xi) = d_1) \text{ if } \theta \in \mathcal{R}_1$$

$$a_1 P_{\theta} (\delta(\xi) = d_0) \leq a_0 P_{\theta} (\delta(\xi) = d_1) \text{ if } \theta \in \mathcal{R}_0$$

(3)

We use that

$$l = P_{\theta}(\delta(X) = d_0) + P_{\theta}(\delta(X) = d_1)$$

to get

$$P_{\theta}(\delta(X) = d_1) \leq \frac{q_1}{q_0 + q_1} \quad \text{if } \theta' \in \Omega_1$$

$$P_{\theta}(\delta(X) = d_1) \geq \frac{q_1}{q_0 + q_1} \quad \text{if } \theta' \in \Omega_0$$

#

Alternative approaches

$$\int R(\theta, \delta) \rho_C$$

$$\int R(\theta, \delta) \lambda(\delta) d\delta$$

or maximum likelihood

$$\max_{\theta \in \Omega} R(\theta, \delta).$$

(4)

Likelihood-based decision rules

Suppose x observation of $\bar{X} \sim F_\theta$, θ unknown. The likelihood is

$$L_x(\theta) = f_\theta(x)$$

where f_θ is the density (pdf or pmf or a mixture). In particular if \bar{X} discrete, it takes a countable number of values $\{x_1, x_2, \dots\}$ with $P_\theta(x) = P_\theta(\bar{X}=x)$, the likelihood is

$$L_x(\theta) = P_\theta(x)$$

We want to base the decision on the value of the likelihood. We want to maximize the likelihood, we talk about gain functions instead of loss functions.

(5)

Suppose there are a countable number of decisions possible $D = \{d_1, d_2, \dots\}$

d_k means $\Theta \in \Omega_k$

for $\Omega = \bigcup_{k=1}^{\infty} \Omega_k$. A gain function is

$g(\Theta, d) = a(\Theta) 1_{\{d\}}$ if d is a correct decision)

with a a positive function. We can weight the likelihood

$L_x(\Theta)$

with $g(\Theta, d)$ (when Θ is true value). We would like to maximize $a(\Theta)L_x(\Theta)$ and select a decision that would be true if the maximizing value would be true value.

For point estimation, one assumes the gain does not depend on θ , to a(θ) is constant, so one maximizes

$$L_x(\theta)$$

which gives us the maximum likelihood estimator.

For hypothesis testing, assume d_0, d_1 are the possible decisions, Ω_0, Ω_1 are the the set of parameter values that define d_0 and d_1 . Assume the gain is g_0 when $\theta \in \Omega_0$ and the decision is correct, and gain is g_1 when $\theta \in \Omega_1$, and the decision is correct.

$$g_0 \sup_{\theta \in \Omega_0} L_x(\theta) > g_1 \sup_{\theta \in \Omega_1} L_x(\theta)$$

we make the decisions d_0 , and if

$$g_0 \sup_{\theta \in \Omega_0} L_x(\theta) < g_1 \sup_{\theta \in \Omega_1} L_x(\theta)$$

(7)

we make the decision d_1 .

Equivalently, make the decision do if

$$\frac{\sup_{\Theta \in \mathcal{R}_0} L_X(\Theta)}{\sup_{\Theta \in \mathcal{R}_1} L_X(\Theta)} > \frac{q_1}{q_0}$$

and make decision d_1 for opposite inequality.

This leads to likelihood ratio (LR) tests.

(Admissible decisions, complete and essentially complete
suffices, is cursory material, read on your own)

Uniformly most powerful tests

$\mathcal{F} = \{F_\theta : \theta \in \mathcal{R}\}$. Partition the

parameter space into $\mathcal{R} = \mathcal{R}_H \cup \mathcal{R}_K$

and $\mathcal{F} = H \cup K = \{F_\theta : \theta \in \mathcal{R}_H\} \cup \{\bar{F}_\theta : \theta \in \mathcal{R}_K\}$

⑧

Possible decisions

$$\begin{array}{ccccc} \text{do } & \text{means} & H & \text{true} \\ d_0 & -/- & K & \text{true} \end{array}$$

and we can let $d_0 = 0, d_1 = 1, D = \{0, 1\}$

We make a test by looking at the data X and the set of possible values that X can take

$$\mathcal{X} = S_0 \cup S_1$$

where

$$S_0 = \{x \in \mathcal{X} : \mathcal{J}(x) = d_0\}$$

$$S_1 = \{x \in \mathcal{X} : \mathcal{J}(x) = d_1\}$$

we call S_1 the critical region, so the acceptance region. Significance level $\alpha \in (0, 1)$ is chosen, and we try to find a critical region S_1 such that

$$(1) \quad P_{\theta}(\mathcal{J}(X) = d_1) = P_{\theta}(X \in S_1) \leq \alpha$$

for ~~every~~ every $\theta \in \Omega_H$.

(9)

subject to this we want to make the power function

$$\beta(\theta) = P_{\theta}(X \in S_i) \quad (2)$$

as large as possible over Θ_K

To find an optimal test, find the critical region S_i which makes (2) as large as possible under the condition (1).

So far we have had deterministic decision rules. We now introduce randomized tests : if x outcome we draw a random ~~rand~~ variable $R(x)$ with two possible outcomes (d_0, d_1) . If we get r_0 we accept it, if we get r_1 we reject it.

(10)

We have

$$P_{\theta}(R(x) = r_1) = \varphi(x)$$

and should only depend on x (it can not depend on θ). The function φ is called the critical function,

$$x \mapsto [0, 1]$$

For each outcome x we

accept H with probability $1 - \varphi(x)$
reject H -/- $\varphi(x)$

Note that a deterministic test is a special case when we have

$$\begin{aligned}\varphi(x) &= 1 & x \in S_1, \\ \varphi(x) &= 0 & x \in S_0.\end{aligned}$$

Now assume we have a randomized test given by φ .

The probability of rejection is

$$\begin{aligned} P_{\theta}(R(\bar{X}) = r_1) &= \int P_{\theta}(R(x) = r_1) dF_{\theta}(x) \\ &= \int \varphi(x) dF_{\theta}(x) \\ &= \underline{\mathbb{E}_{\theta}(\varphi(\bar{X}))} \end{aligned}$$

Finding an optimal test then means
maximize

$$\beta_{\varphi}(\theta) = \underline{\mathbb{E}_{\theta} \varphi(\bar{X})}$$

when $\theta \in \mathcal{R}_K$, subject to

$$\underline{\mathbb{E}_{\theta} \varphi(\bar{X})} \leq \alpha$$

when $\theta \in \mathcal{R}_H$ (for all θ in \mathcal{R}_H).

(12)