STAT3602 Statistical Inference
Example Class 1

LIU Chen

Department of Statistics and Actuarial Science, HKU

07/09/2020

Outline

Parameter space Θ

- collection of k unknown parameters, usually a subset of $\mathbb{R}^k$.
- The true parameter is some **unknown** $\theta \in \Theta$

Sample space S

collection of all possible realisations x of a random vector X.

Action space $\mathcal{A}$

collection of all possible actions under consideration

Review: Decision Problem: Frequentist Approach

Statistical Model

- a link between Θ and S
- a family of probability functions $\{f(\cdot|\theta) : \theta \in \Theta\}$

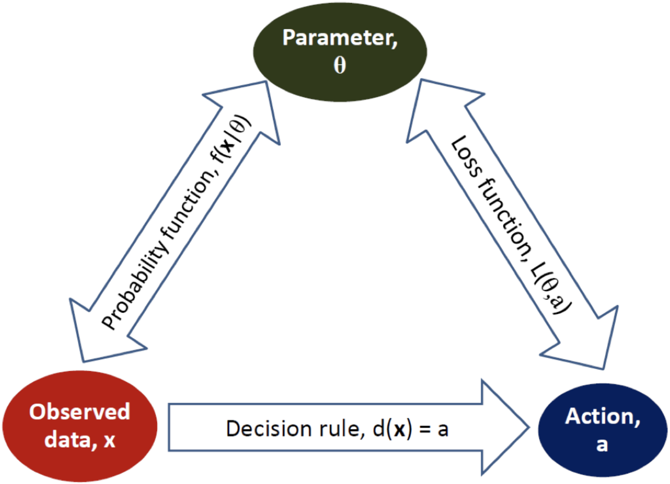
Loss function $L : \Theta \times \mathcal{A} \to \mathbb{R}$

$L(\theta, a)$ is the *loss* incurred by taking action a when θ is the true parameter

Decision rule $d : S \to A$

an action, $d(x)$, to be taken when X is observed to be $x \in S$

Review: Formulation



The risk function of the decision rule $d(\cdot)$ is the expected loss incurred by adopting decision rule d under each possible $\theta \in \Theta$, i.e.

$$R(\theta, d) = \mathbb{E}_\theta [L(\theta, d(x))] \tag{1}$$

- if X is continuous,

$$R(\theta, d) = \int_S L(\theta, d(x)) f(x|\theta) \, dx; \tag{2}$$

- if X is discrete,

$$R(\theta, d) = \sum_{x \in S} L(\theta, d(x)) f(x|\theta). \tag{3}$$

Outline

Exercise 1: Background

Each semester evening between Sunday and Thursday, the HKU superintendent has to decide whether to call off the next day school because of Tropical Storm/Black Rainstorm conditions.
If he fails to call off school and there is storm, there are various possible consequences, including students and teachers failing to show up for school, the possibility of traffic accidents etc.
If he calls off school, then regardless of whether there actually is storm that day there will have to be a make-up day later in the year. After weighing up all the possible outcomes he decides that the costs of failing to close school when there is storm are twice the costs incurred by closing school. If he does not call off school and there is no storm, then of course there is no loss.

Exercise 1: Background

Two local radio stations give independent and identically distributed weather forecasts. If there is to be storm, each station will forecast this with probability $\frac{3}{4}$, but predict no storm with probability $\frac{1}{4}$. If there is to be no storm, each station predicts storm with probability $\frac{1}{2}$.
The superintendent will listen to the two forecasts this evening, and then make his decision on the basis of the number of stations forecasting storm.

Formulate the above into a decision problem and calculate the risk functions corresponding to the deterministic rules.

Exercise 1: Formulation

Parameter space:

$$\Theta = \{0, 1\}, \tag{4}$$

where "this is no storm" is coded as $\theta = 0$ and "this is storm" as $\theta = 1$

Sample space:

$$S = \{0, 1, 2\}, \tag{5}$$

where $x \in S$ refers to the number of stations forcasting the storm.

Action space:

$$\mathcal{A} = \{0, 1\}, \tag{6}$$

where "not call off" is coded as $a = 0$ and "call off" as $a = 1$

Exercise 1: Statistical Model $f(x|\theta)$

$f(x\|\theta)$		x: 0	x: 1	x: 2
θ	0			
	1			

Exercise 1: Statistical Model $f(x|\theta)$

$f(x\|\theta)$		x = 0	x = 1	x = 2
θ	0	$1/4$	$1/2$	$1/4$
	1	$1/16$	$3/8$	$9/16$

Exercise 1: Loss Function

		θ	
$L(\theta, a)$		0	1
a	0		
	1		

Exercise 1: Loss Function

		θ	
$L(\theta, a)$		0	1
a	0	0	2
	1	1	1

Exercise 1: Decision Rule

$d(x)$		d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8
	0	0	0	0	0	1	1	1	1
x	1	0	0	1	1	0	0	1	1
	2	0	1	0	1	0	1	0	1

Exercise 1: Risk Functions

$R(\theta, d)$		d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8
θ	0								
	1								

$R_1 = E[L(\theta, d_1(x))] = L(\theta, d_1(0))P_0 + L(\theta, d_1(1))P_1 + L(\theta, d_1(2))P_2$

$\theta = 0,$

$$\begin{aligned} R_1 &= L(0, d_1(0))1/4 + L(0, d_1(1))1/2 + L(0, d_1(2))1/4 \\ &= L(0,0)1/4 + L(0,0)1/2 + L(0,0)1/4 = 0 \end{aligned}$$

$\theta = 1,$

$$\begin{aligned} R_1 &= L(1, d_1(0))1/16 + L(1, d_1(1))3/8 + L(1, d_1(2))9/16 = 2 \\ &= L(1,0) = 2 \end{aligned}$$

Exercise 1: Risk Functions

$$R_1 = E[L(\theta, d_1(x))] = L(\theta, d_1(0))P_0 + L(\theta, d_1(1))P_1 + L(\theta, d_1(2))P_2$$

$\theta = 0,$

$$\begin{aligned} R_1 &= L(0, d_1(0))1/4 + L(0, d_1(1))1/2 + L(0, d_1(2))1/4 \\ &= L(0,0)1/4 + L(0,0)1/2 + L(0,0)1/4 = 0 \end{aligned}$$

$\theta = 1,$

$$\begin{aligned} R_1 &= L(1, d_1(0))1/16 + L(1, d_1(1))3/8 + L(1, d_1(2))9/16 = 2 \\ &= L(1,0) = 2 \end{aligned}$$

$R(\theta, d)$		d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8
θ	0	0	1/4	1/2	3/4	1/4	1/2	3/4	1
	1	2	23/16	13/8	17/16	31/16	11/8	25/16	1

Review: Admissibility

Definition. A rule d *strictly dominates* another rule d^* if

$$R(\theta, d) \leq R(\theta, d^*) \text{ for all } \theta \in \Theta \quad \text{and} \quad R(\theta', d) < R(\theta', d^*) \text{ for some } \theta' \in \Theta.$$

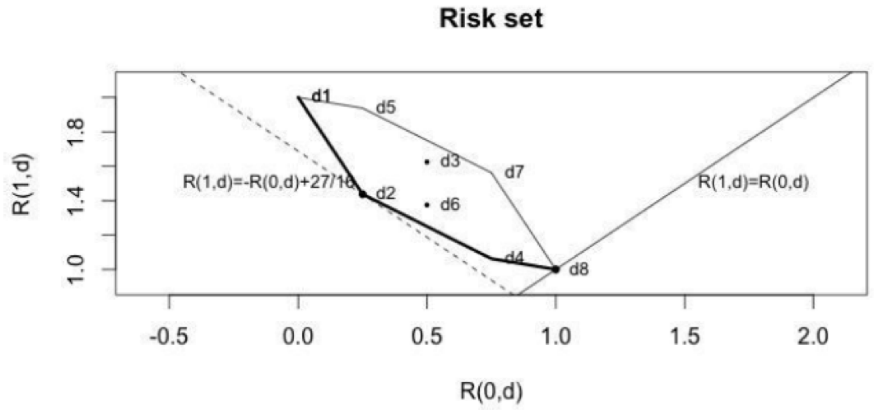
If d strictly dominates d^*, then obviously d is the better choice.

Definition. A rule strictly dominated by another rule is *inadmissible*.

Definition. If d is not inadmissible, then it is said to be *admissible*.

Excise 1: Admissibility. Admissible: d_1, d_2, d_4, d_8

$R(\theta, d)$		d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8
θ	0	0	1/4	1/2	3/4	1/4	1/2	3/4	1
	1	2	23/16	13/8	17/16	31/16	11/8	25/16	1



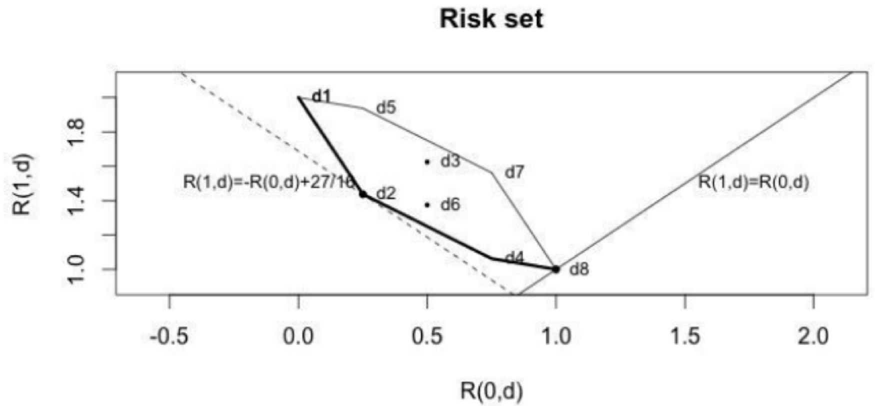
Review: Minimaxity

Definition. A rule d is *minimax* if, for all possible rules d',

$$\sup\{R(\theta, d') : \theta \in \Theta\} \geq \sup\{R(\theta, d) : \theta \in \Theta\}.$$

Excise 1: Minimaxity. d_8 is *minimax*

$R(\theta, d)$		d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8
θ	0	0	1/4	1/2	3/4	1/4	1/2	3/4	1
	1	2	23/16	13/8	17/16	31/16	11/8	25/16	1



Outline

Exercise 2

Let X be uniformly distributed on $[0, \theta]$, where $\theta \in (\theta, +\infty)$ is an unknown parameter. Let the action space be $[0, +\infty)$ and the loss function $L(\theta, a) = (\theta - a)^2$, where a is the action chosen. Consider the decision rules $d_\mu(x) = \mu x$, $\mu \geq 0$.

For what value of μ is d_μ unbiased?

Review: Unbiased rule

Definition. A rule d is *unbiased* if

$$\boxed{\underbrace{\mathbb{E}_\theta L(\theta', d(\boldsymbol{X}))}_{\text{(expected loss w.r.t. false } \theta\text{)}} \geq \underbrace{R(\theta, d)}_{\text{(expected loss w.r.t. true } \theta\text{)}} \quad \forall\, \theta, \theta' \in \Theta}$$

Exercise 2: Solution

If d_μ is unbiased, then we could have

$$R(\theta, d_\mu) \leq \mathbb{E}_\theta \left[L\left(\theta', d_\mu(X)\right) \right], \forall \theta' \in (0, +\infty) \tag{7}$$

And the following equality holds,

$$R(\theta, d_\mu) = \mathbb{E}_\theta \left[L\left(\theta, d_\mu(X)\right) \right] = \left(1 - \mu + \frac{1}{3}\mu^2\right)\theta^2; \tag{8}$$

$$\mathbb{E}\left[L\left(\theta', d_\mu(X)\right) \right] = \int_0^\theta \left(\theta' - \mu x\right)^2 \theta^{-1} dx = \theta'^2 - \mu\theta'\theta + \frac{1}{3}\mu^2\theta^2. \tag{9}$$

Thus, the following inequality holds,

$$\left(1 - \mu + \frac{1}{3}\mu^2\right)\theta^2 \geq \theta'^2 - \mu\theta'\theta + \frac{1}{3}\mu^2\theta^2, \forall \theta, \theta' \in (0, +\infty). \tag{10}$$

Exercise 2: Solution (CONTI.)

Thus, we could have

$$\left(\theta' - \frac{1}{2}\mu\theta\right)^2 + (\mu - 1)\,\theta^2 - \frac{\mu^2\theta^2}{4} \geq 0, \forall\theta, \theta' \in (0, +\infty)\,, \forall\theta, \theta' \in (0, +\infty)\,. \tag{11}$$

Thus, $\mu = 2$

Practice more!

Past paper:

- 2014. Question 1, (c) (just use the result of (b))
- 2013. Question 1, (a) – (c)