Introduction to probability

Statistics is concerned with **decision making** in the face of **uncertainty**.

Two kinds of uncertainty:

- Uncertainty due to randomness.
- Uncertainty due to uncertainty of state of nature.

We can estimate the state of nature through observations. The sample size depends on the cost of making an observation and the cost of making the wrong decision.

Basic framework

Decisions involve three components– **actions, states of nature** and **outcomes.**

	Explosive states	Nonexplosive states
Light a match	Explosion	Brightness, No explosion
Not light a match	No explosion, Darkness	No explosion, Darkness

	Quiz given (0.25)	Quiz not given(0.75)
Go to class	Take quiz, miss game(-4)	Miss game(6)
Stay at home	Miss quiz, enjoy game(4)	Enjoy game(-6)

Q: which action is best?

It is hard to find an action that is best over all states of nature.

- minimax principle: pick an action for which the maximum loss is minimized. Rule for pessimists.
- Minimize expected loss: pick an action for which the expected loss is minimized.

Minimizing expected loss

	Smith dies(0.05)	Smith lives(0.95)
Insures Smith	25,000	-100
Not Insure Smith	0	0

Game and decision problem

Two contestants, Statistician and Nature, put up 1 or 2 fingers simultaneously. Statistician wins if sum is even. The numbers are the loss to the statistician.

	Nature 1	Nature 2
Statistician 1	-2	3
Statistician 2	3	-4

In summary, the basic elements of a decision problem are

- A nonempty set, Θ, of possible states of nature.
- A nonempty set, $\mathcal{A}$, of actions available.
- A loss function, $L(\theta, a)$, a real valued function on $\Theta \times \mathcal{A}$.

Decision function, Risk function

A statistical decision problem is a game $(\Theta, \mathcal{A}, \mathcal{L})$ coupled with an experiment involving a random variable X whose distribution P_θ depend on the state $\theta \in \Theta$. On the basis of x, the statistician chooses an action $d(x) \in \mathcal{A}$. The loss now is $L(\theta, d(X))$.

The expected value of $L(\theta, d(X))$ is called the **risk function**.
$R(\theta, d) = E_\theta L(\theta, d(X))$.

Example

Losses incurred by contractor

	θ_1: Use 15 amp	θ_2: Use 20 amp	θ_3: Use 30 amp
Install 15 amp	1	5	7
Install 20 amp	2	2	6
Install 30 amp	3	3	3

Experiment

Ask the residents how many amperes they use.

	θ_1: Use 15 amp	θ_2: Use 20 amp	θ_3: Use 30 amp
$x_1 = 10$amp	$1/2$	0	0
$x_2 = 12$ amp	$1/2$	$1/2$	0
$x_3 = 15$ amp	0	$1/2$	$1/3$
$x_4 = 20$ amp	0	0	$2/3$

Strategy

A mapping from X to $\mathcal{A}$: Define which action to take for each possible value of the observation(s).

	$x_1 = 10amp$	$x_2 = 12amp$	$x_3 = 15amp$	$x_4 = 20amp$
strategy s_1	a1	a1	a2	a3
s_2	a1	a2	a3	a3
s_3	a3	a3	a3	a3
s_4	a1	a1	a1	a1
s_5	a3	a3	a2	a1

Risk

	θ_1: Use 15 amp	θ_2: Use 20 amp	θ_3: Use 30 amp
s_1	1	3.5	4
s_2	1.5	2.5	3
s_3	3	3	3
s_4	1	5	7
s_5	3	2.5	6.67

s_1 dominates s_4. Any other dominated strategies?

Framework of decision making

1. $\mathcal{A}$=$(a_1, a_2, \cdots)$, a set of available actions.
2. $\Theta = (\theta_1, \theta_2, \cdots)$, a set of possible states of nature.
3. Loss function or loss table $L(\theta, a)$.
4. Observations from experiments $X = (x_1, x_2, \cdots)$.
5. The probability distribution of X, $P_\theta(x)$.
6. Possible strategies $S = (s_1, s_2, \cdots)$ which is mapping from x to a: $d(x) \in \mathcal{A}$.
7. Risk function $R(\theta, d) = E_\theta L(\theta, d(X))$.

exercise

Jane Smith can cook spaghetti, hamburger, or steak for dinner. She has learned from past experience that if her husband is in a good mood she can serve him spaghetti and save money, but if he is in bad mood, only a juicy steak will calm him down and make him bearable. In short, there are three actions:

a_1 : prepare spaghetti

a_2 : prepare hamburger

a_3 : prepare steak.

and three states of nature:

θ_1 : Mr. Smith is in a good mood

θ_2 : Mr. Smith is in a normal mood.

θ_3 : Mr. Smith is in a bad mood.

The loss table is

	θ_1	θ_2	θ_3
a_1	0	5	10
a_2	2	3	9
a_3	4	5	6

The experiment she performs is to tell him when he returns home that she lost the afternoon paper. She foresees 4 possible responses. They are

x_1 : "Newspapers will get lost".

x_2 : "I keep telling you a place for everything and everything in its place".

x_3 : "Why did I ever get married?"

x_4 : an absent-minded, far-away look.

The distribution of the observation is

	x_1	x_2	x_3	x_4
θ_1	0.5	0.4	0.1	0
θ_2	0.2	0.5	0.2	0.1
θ_3	0	0.2	0.5	0.3

List 4 strategies and evaluate their risks.
Find the minimax risk strategy.
Which is the best strategy if it is known that Mr. Smith is in good mood 30% of the time and in normal mood 50% of the time?

Odd or even finger problem. $X =$ number of fingers nature tells statistician he will put up.

$\theta = 1$: $P(X = 1) = 3/4, P(X = 2) = 1/4.$

$\theta = 2$: $P(X = 2) = 3/4, P(X = 1) = 1/4.$

Four decision rules:

$s_1 : s_1(x = 1) = 1, s_1(x = 2) = 1.$

$s_2 : s_2(x = 1) = 1, s_2(x = 2) = 2$

$s_3 : s_3(x = 1) = 2, s_3(x = 2) = 1$

$s_4 : s_4(x = 1) = 2, s_4(x = 2) = 2.$

	s_1	s_2	s_3	s_4
$\theta_1 = 1$	-2	-3/4	7/4	3
$\theta_2 = 2$	3	-9/4	5/4	-4

Loss table

	a_1	a_2	a_3	a_4
θ_1	0	2	4	8
θ_2	6	4	2	5
θ_3	3	2	1	0

Distribution of observations given state of nature

	z_1	z_2	z_3	z_4
$\theta_1(0.2)$	0	0.4	0.6	0
$\theta_2(0.3)$	0.4	0.2	0.2	0.2
$\theta_3(0.5)$	0.1	0.2	0.3	0.4

```
Actions by Certain Strategies
                 Response
Strategies      z1   z2   z3   z4
   s1           a1   a2   a3   a4
   s2           a2   a3   a4   a3
   s3           a3   a3   a3   a3
   s4           a4   a4   a1   a1
```

a). How many possible strategies are there?
b). Evaluate the average losses for strategies s1, s2, s3 and s4.
c). Among these 4 strategies listed above, find the minimax strategy
d). If it is known the 3 of nature are equally likely, find the strategy that minimizes the average risk.

Find the best strategy

How to find the strategy with minimum average risk (or Bayes risk) given the prior distribution on the states of nature?
Only need to find the best strategy (minimizes the conditional expected loss) given each observation X.
In the ampere example, given $X = x_1$, find $P(\theta|x_1)$ and take action which minimizes the loss averaged over the distribution of $(\theta|x_1)$.
Repeat for $X = x_i, i = 2, 3, 4$.