STA 371G: Statistics and Modeling

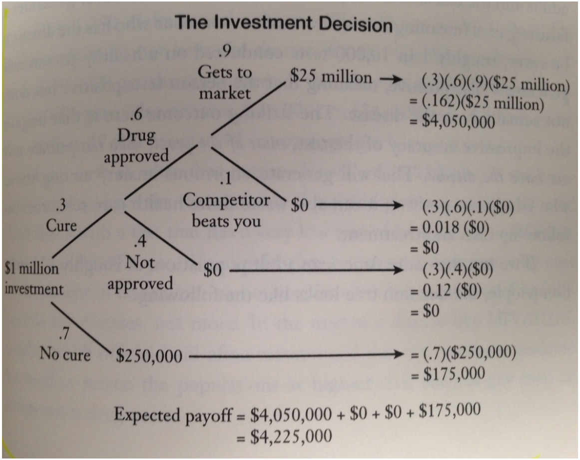
Decision Making Under Uncertainty: Decision Trees

Mingyuan Zhou
McCombs School of Business
The University of Texas at Austin

http://mingyuanzhou.github.io/STA371G

Probability and Decision Making

Suppose you are presented with an investment opportunity in the development of a drug... probabilities are a vehicle to help us build scenarios and make decisions.



Examples of Decision Making Under Uncertainty

- Contract bidding
- Marketing strategy
- When to introduce a new car model
- Inventory management
- Insurance
- Lending
- Investment
- Oil drilling
- Skip a midterm test and add its weight to the final
- Express or ground shipping

Motivating Example

The company SciTools has been invited to make a bid on a government contract to supply a set of scientific instruments.

- SciTools estimates that working on the proposal will cost \$5,000 and \$95,000 to supply the instruments
- SciTools believes that there is a 30% chance that there will be no competing bid. Based on previous bidding processes, SciTools believes in the following set of probabilities describing the possible competing (lowest) bids:

Bid	Probability
$< \$115,000$	0.2
Between \$115,000 and \$120,000	0.4
Between \$120,000 and \$125,000	0.3
$> \$125,000$	0.1

How much should SciTools bid?

Payoff Table:

Probability:	0.3	0.14	0.28	0.21	0.07	
Competitor:	No bid	<115	(115,120)	(120,125)	>125	ER
No bid						
Bid 115						
Bid 120						
Bid 125						

Loss Table:

Probability:	0.3	0.14	0.28	0.21	0.07	
Competitor:	No bid	<115	(115,120)	(120,125)	>125	EL
No bid						
Bid 115						
Bid 120						
Bid 125						

How much to bid?
minimax:
maxmin:
minimaxloss:
ER or *EL*:

Elements of Decision Analysis

1. Payoff Tables: list the payoff for each possible outcome based on a decision. In general we think about this in monetary values but it could be more generally defined (think about the concept of utility).
2. Decision Criteria: in order to define a “best” decision, we need to work with a decision criteria. The most used criteria is the maximization of Expected Payoff or Expected Utility.
3. Expected Payoff (ER): or Expected Monetary Value (EMV), is the weighted average of the possible payoffs for a decision, weighted by the corresponding probabilities.

Payoff Tables

- We are faced with 3 possible decisions: D_1, D_2 and D_3
- There are 3 possible outcomes: O_1, O_2 and O_3
- The payoff table is defined as:

	O_1	O_2	O_3
D_1	10	10	10
D_2	-10	20	40
D_3	-30	30	70

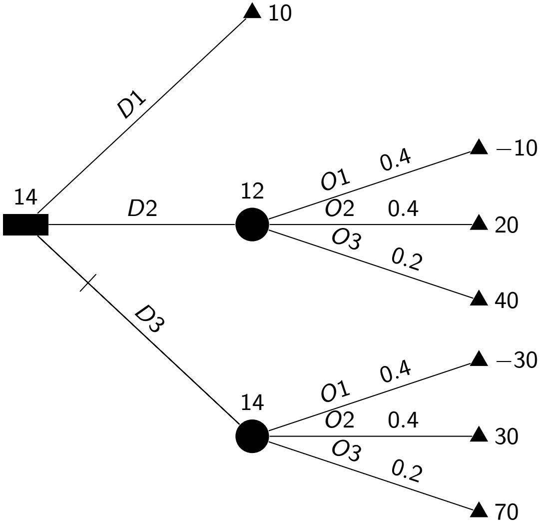
- The probabilities of each outcome are $p(O_1) = 0.4$, $p(O_2) = 0.4$ and $p(O_3) = 0.2$

Which decision should you choose?

Decision Trees

- A decision tree is a graphical tool to represent a decision problem
- The tree is composed by nodes (circles, squares and triangles) and branches (lines)
- Three types of nodes: decision node (square), probability node (circle) and end node (triangle)
- Time proceeds from left to right...

Decision Trees

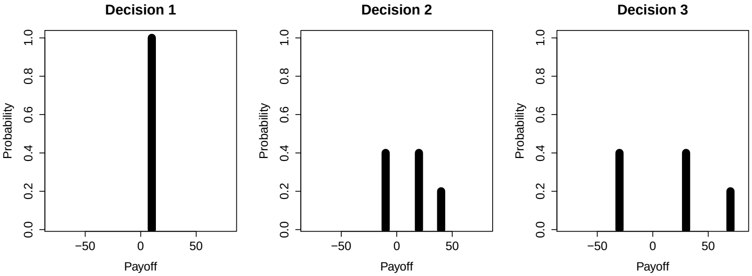


Folding-Back Procedure

- Start from the right side of the decision tree and work back to the left
    1. at each probability node, calculate the Expected Payoff (ER)
    2. at each decision node, take a maximum of all possible ER's to identify the optimal decision

Risk Profiles

The *risk profile* for a decision is a bar chart that represents the probability distribution of payoffs for this decision.



Sensitivity Analysis

The optimal decision that maximizes the expected payoff may change as the input variables (such as the probabilities for different states) change. Sensitivity analysis will help examine these relationships.

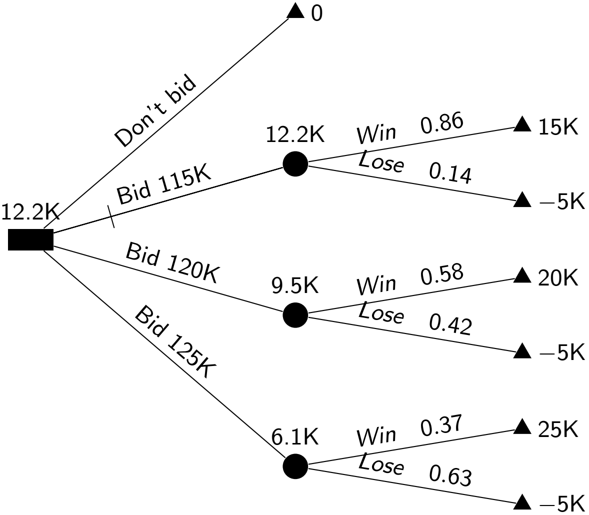
For example, sensitivity analysis can be performed to find whether a particular probability is important in determining the preferred action, and it can be very useful in situations where the value of that probability is not known precisely.

Example

Payoff Table:

Probability:	0.3	0.14	0.28	0.21	0.07	
Competitor:	No bid	$<$115	(115,120)	(120,125)	$>$125	ER
No bid	0	0	0	0	0	0
Bid 115	15	-5	15	15	15	12.2
Bid 120	20	-5	-5	20	20	9.5
Bid 125	25	-5	-5	-5	25	6.1

Decision Trees



Decision Making and Bayes' Theorem

Example:

Our goal is to decide whether to test an athlete for drugs.

We are uncertain about whether an athlete uses drugs or not
($D = 1$ or 0)

and whether the athlete would test positive or not
($T = 1$ or 0).

Decision Making and Bayes' Theorem

We have to make an overall decision about whether to test or not. We'll denote this decision by $S = 1$ (test) and $S = 0$ (don't test). S is for *screen*.

We also have to make a decision about whether to ban the athlete. We will denote this decision by $B = 1$ or 0, for ban and not ban.

Let's assume that if we do not test, we do not ban.

What are our costs and benefits?

Example: Testing College Athletes

- **Goal:** We need to decide whether or not to test all college athletes for the use of performance enhancing drugs. If an athlete is found guilty we would have to ban them from college sports.
- What are the benefits? How about the costs? Let's assume the following
  1. benefit G for correctly identifying a cheater
  2. cost C_1 for testing each athlete
  3. cost C_2 for falsely accusing a non-user
  4. cost C_3 not identifying a drug user
  5. cost C_4 for violating non-users' privacy

We will use $G = 25$, $C_1 = 1$, $C_2 = 50$, $C_3 = 20$, and $C_4 = 2$.
If you screen, this gives payouts:

- ban a drug user: $G - C_1 = 24$
- ban a non-user (ouch!): $-C_1 - C_2 - C_4 =$ -53
- don't ban a drug user: $-C_1 - C_3 =$ -21
- don't ban a non-user: $-C_1 - C_4 =$ -3

If you don't screen, this gives payouts

- don't ban a drug user: $-C_3 =$ -20
- don't ban a non-user: 0

What about T and D?

There is a 5% chance the athlete is a user.

$$P(D = 1) = .05$$

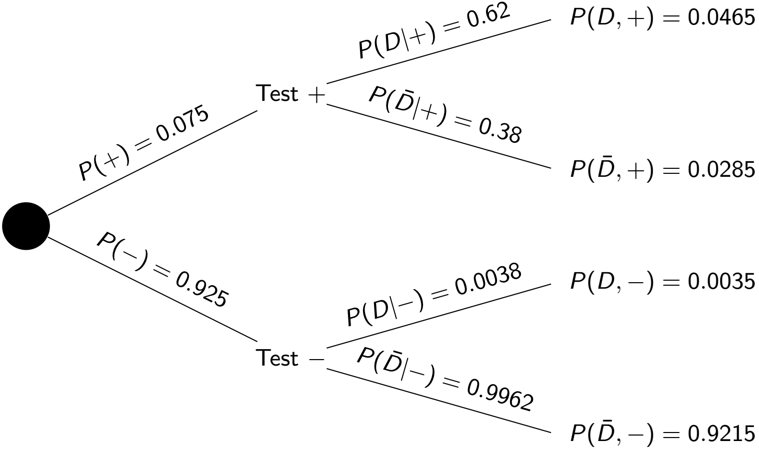
There is

- a 3% chance of a false positive
- a 7% chance of a false negative.

$$P(T = 1 \mid D = 0) = .03, \quad P(T = 0 \mid D = 1) = .07.$$

This gives the joint probability table and tree diagram

	D	$\bar{D}$	
$+$	0.0465	0.0285	0.075
$-$	0.0035	0.9215	0.925
	0.05	0.95	1



Example: Testing College Athletes

The **payoff table** if NO test is:

Decision	D	$\bar{D}$
Bar	B	$-C_2$
Don't	$-C_3$	0

The **payoff table** for TESTING is:

	(D and $+$)	($\bar{D}$ and $+$)	(D and $-$)	($\bar{D}$ and $-$)
Bar	$B - C_1$	$-(C_1 + C_2 + C_4)$	$B - C_1$	$-(C_1 + C_2 + C_4)$
Don't	$-(C_1 + C_3)$	$-(C_1 + C_4)$	$-(C_1 + C_3)$	$-(C_1 + C_4)$

Lets further assume that $B = 25,\ C_1 = 1,\ C_2 = 50,\ C_3 = 20$ and $C_4 = 2$... Your job is to build a decision tree to decide whether or not to conduct the test for all athletes

Example: Testing College Athletes

- The **payoff table** if NO test is:

Decision	D	$\bar{D}$
Bar	25	-50
Don't	-20	0

- Expected Payoff for NO test:

$$\begin{aligned} ER|NO\ \ Test &= -20 * P(D) + 0 * P(\bar{D}) \\ &= 0.05 * -20 = -1 \end{aligned}$$

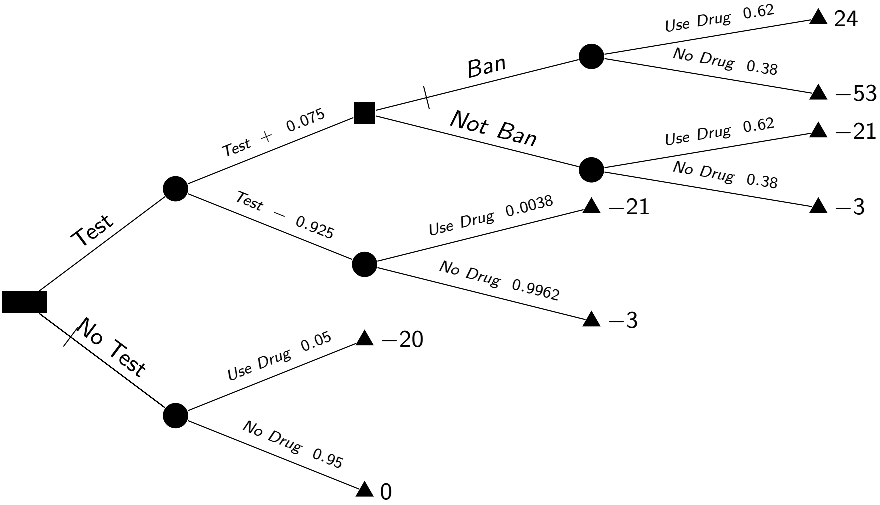
- The **payoff table** for TESTING is:

	(D and $+$)	($\bar{D}$ and $+$)	(D and $-$)	($\bar{D}$ and $-$)
Bar	24	-53	24	-53
Don't	-21	-3	-21	-3

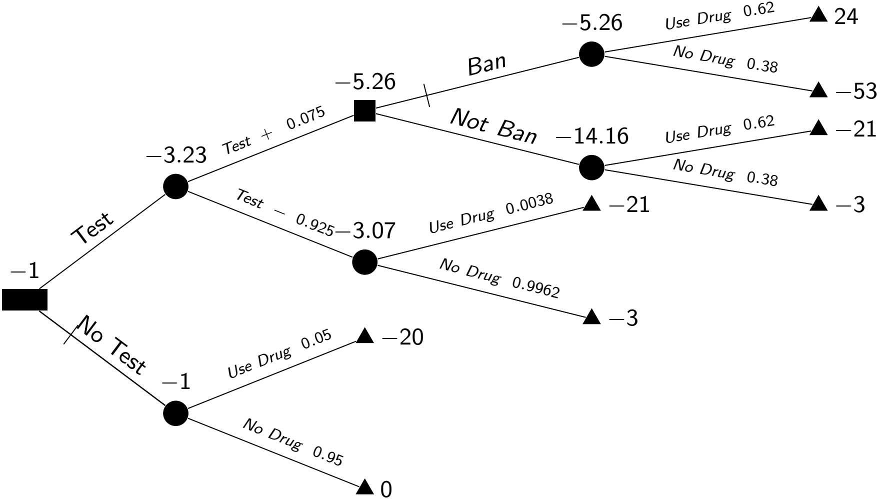
- Expected Payoff for Testing (Bar if tested positive):

$$\begin{aligned} ER|Test &= 24 * P(D, +) - 53 * P(\bar{D}, +) \\ &\quad + 21 * P(D, -) - 3 * P(\bar{D}, -) \\ &= -3.2325 \end{aligned}$$

Decision Trees

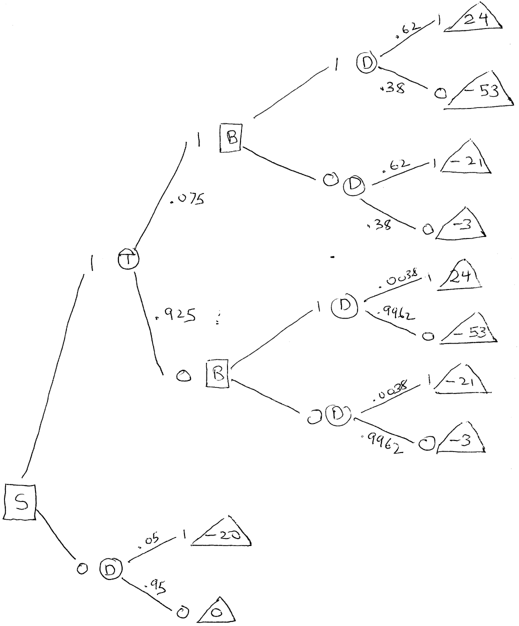


Decision Trees



All the outcomes, decisions, and payoffs.

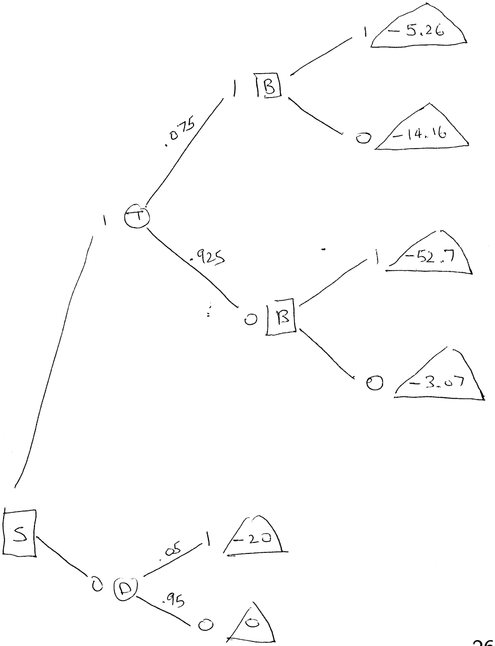
For example, if $B = 1$ and $D = 1$, then you banned a cheater, and the payoff is $G - C_1 = 25 - 1 = 24$; if $B = 1$ and $D = 0$, then you banned an innocent, and the payoff is -53.



Average out the D outcomes.

$0.62*24+0.38*(-53)$
$= -5.26$

$0.0038 * (-21) + 0.9962 * (-3)$
$= -3.07$



Average out the T outcomes.

.075*(-5.26) + .925*(-3.07)
= -3.23425

It is best not to screen!!!