6.01: Introduction to EECS I

Discrete Probability and State Estimation

April 12, 2011

Midterm Examination #2

Time: Tonight, April 12, 7:30 PM to 9:30 PM

Location: Walker Memorial (if last name starts with A-M)
10-250 (if last name starts with N-Z)

Coverage: Everything up to and including Design Lab 9.

You may refer to any printed materials that you bring to exam.

You may use a calculator.

You may not use a computer, phone, or music player.

No software lab or design lab this week. Instead, there are extensive tutor problems (week 10). Extra office hours Thursday and Friday, 9:30am-12:30pm and 2pm-5pm.

6.01: Overview and Perspective

The **intellectual themes** in 6.01 are recurring themes in EECS:

- design of complex systems
- modeling and controlling physical systems
- augmenting physical systems with computation
- building systems that are robust to uncertainty

Intellectual themes are developed in context of a mobile robot.

Goal is to convey a distinct perspective about engineering.

Module 1: Software Engineering

Focus on abstraction and modularity.

Topics: procedures, data structures, objects, state machines

Lab Exercises: implementing robot controllers as state machines

SensorInput → Brain → Action

Abstraction and Modularity: Combinators

`Cascade`: make new SM by cascading two SM's
`Parallel`: make new SM by running two SM's in parallel
`Select`: combine two inputs to get one output

Themes: PCAP

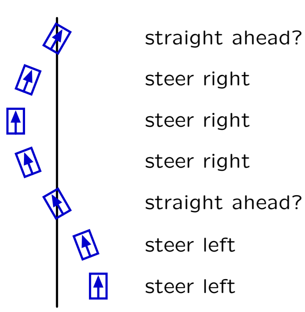
Primitives – **C**ombination – **A**bstraction – **P**atterns

Module 2: Signals and Systems

Focus on discrete-time feedback and control.

Topics: difference equations, system functions, controllers.

Lab exercises: robotic steering



Themes: modeling complex systems, analyzing behaviors

Module 3: Circuits

Focus on resistive networks and op amps.

Topics: KVL, KCL, Op-Amps, Thevenin equivalents.

Lab Exercises: build robot "head":

- motor servo controller (rotating "neck")
- phototransistor (robot "eyes")
- integrate to make a light tracking system

Themes: design and analysis of physical systems

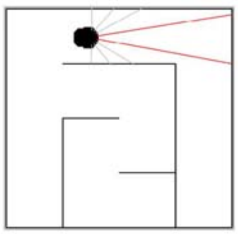
Module 4: Probability and Planning

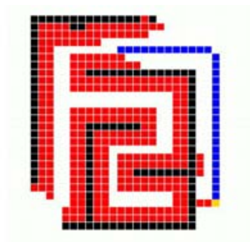
Modeling uncertainty and making robust plans.

Topics: Bayes' theorem, search strategies

Lab exercises:

- Mapping: drive robot around unknown space and make map.
- Localization: give robot map and ask it to find where it is.
- Planning: plot a route to a goal in a maze





Themes: Robust design in the face of uncertainty

Probability Theory

We will begin with a brief introduction to probability theory.

Probability theory provides a framework for

- reasoning about uncertainty
  - making precise statements about uncertain situations
  - drawing reliable inferences from unreliable observations
- designing systems that are robust and fault-tolerant

Let's Make a Deal

The game:

- There are four lego bricks in a bag.
- The lego bricks are either white or red.
- You get to pull one lego brick out of the bag.
- I give you $\begin{cases} \$20 & \text{if the brick is red} \\ \$0 & \text{otherwise} \end{cases}$

How much would you pay to play this game?

Events

Probabilities are assigned to **events**, which are possible outcomes of an experiment.

Example: flip three coins in succession — possible events:

- head, head, head
- head, tail, head
- one head and two tails
- first toss was a head

There are eight **atomic** (finest grain) events:
HHH, HHT, HTH, HTT, THH, THT, TTH, TTT.

Atomic events are **mutually exclusive** (only one can happen).

Set of all atomic events is **collectively exhaustive** (cover all cases).

Set of all possible atomic events is called the **sample space** U.

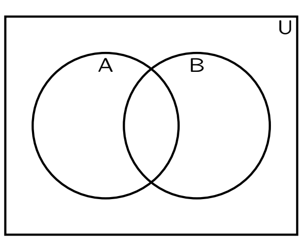
Axioms of Probability

Probability theory derives from three axioms:

- **non-negativity:** $\Pr(A) \geq 0$ for all events A
- **scaling:** $\Pr(U) = 1$
- **additivity:** if $A \cap B$ is empty, $\Pr(A \cup B) = \Pr(A) + \Pr(B)$

From these three, it is easy to prove many other relations.

Example: $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$



Check Yourself

Experiment: roll a fair six-sided die.

Find probability that result is odd and greater than 3.

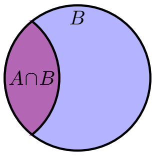
1. $1/6$
2. $2/6$
3. $3/6$
4. $4/6$
5. 0

Conditional Probability

Bayes' rule specifies the probability that one event (A) occurs given that a different event (B) is known to have occurred.

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

Conditioning (on B) restricts the sample space (which was U) to B.



Check Yourself

What is the conditional probability of getting a die roll greater than 3, given that it is odd?

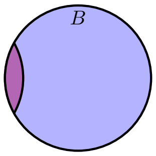
1. $1/2$
2. $1/3$
3. $1/4$
4. $1/5$
5. none of the above

Conditional Probability

Conditioning can increase or decrease the probability of an event.

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

Conditioning can **decrease** the probability of an event.

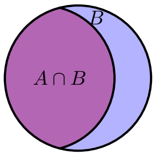


Conditional Probability

Conditioning can increase or decrease the probability of an event.

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

Conditioning can **increase** the probability of an event.



Random variables

A random variable is the probabilistic analog of a (deterministic) variable.

While the value of a deterministic variable is a number, the value of a random variable is drawn from a **distribution**.

Example: Let X represent the result of the toss of a die.

Then X can take on one of six possible values from a distribution:

event	probability
$X = 1$	$1/6$
$X = 2$	$1/6$
$X = 3$	$1/6$
$X = 4$	$1/6$
$X = 5$	$1/6$
$X = 6$	$1/6$

Random variables

Using random variables can simplify our notation.

Pr($X = 3$) replaces Pr(result of toss is three)

This is especially useful when the sample space is multi-dimensional.

Joint Probability Distributions

Probability laws for multi-dimensional sample spaces are given by joint probability distributions.

Let V represent the toss of the first die and W represent the toss of the second die.

- Pr(V,W) represents the joint probability distribution.
- Pr(v,w) represents the Pr$(V=v$ **and** $W=w)$.

Reducing Dimensionality

The dimensionality of a joint probability distribution can be reduced in two very different ways:

Marginalizing refers to collapsing one or more dimensions by summing over all possible outcomes along those dimensions.

— sum along the collapsed dimension(s)

Conditioning refers to collapsing dimensions by accounting for new information that restricts outcomes.

— apply Bayes' rule

Reducing Dimensionality

Example: prevalence and testing for AIDS.

Consider the effectiveness of a test for AIDS.

We divide the population along two dimensions:
– patients with or without AIDS
– patients for with the TEST is positive or negative

We organize data as a joint probability distribution:

	AIDS	
TEST	true	false
positive	0.003648	0.022915
negative	0.000052	0.973385

How effective is the test?

What is the probability that the test is positive given that the subject has AIDS?

	AIDS	
TEST	true	false
positive	0.003648	0.022915
negative	0.000052	0.973385

1. > 90%
2. between 50 and 90%
3. < 50%
4. cannot tell from this data

How effective is the test?

What is the probability that a subject has AIDS given the TEST is positive?

	AIDS	
TEST	true	false
positive	0.003648	0.022915
negative	0.000052	0.973385

1. > 90%
2. between 50 and 90%
3. < 50%
4. cannot tell from this data

How effective is the test?

Q: Why are previous conditional probabilities so different?
A: Because **marginal** probability of having AIDS is small.

	AIDS	
TEST	true	false
positive	0.003648	0.022915
negative	0.000052	0.973385
	0.003700	0.996300

DDist class

Probability distributions are represented as instances of the **DDist** (discrete distribution) class.

```
class DDist:
    def __init__(self, dictionary):
        self.d = dictionary
    def prob(self, elt):
        if elt in self.d:
            return self.d[elt]
        else:
            return 0
```

Instances are created from Python **dictionaries** that associate atomic events (keys) with probabilities (values).

DDist Example

Example: discrete distribution for toss of a fair coin

```
>>> import lib601.dist as dist
>>> toss = dist.DDist('head':0.5, 'tail':0.5)
>>> toss.prob('head')
0.5
>>> toss.prob('tail')
0.5
>>> toss.prob('H')
0
```

Notice that undefined events return probability 0.

Conditional Distributions

Conditional distributions are represented as procedures.

```
def TESTgivenAIDS(AIDS):
    if AIDS == 'true':
        return dist.DDist({'positive':0.985946,'negative':0.014054})
    else:
        return dist.DDist({'positive':0.023000,'negative':0.977000})


>>> TESTgivenAIDS('true')
DDist('positive':0.985946,'negative':0.014054)
>>> TESTgivenAIDS('true').prob('negative')
0.014054
```

Joint Probability Distributions

Joint probability distributions are represented as discrete distributions with keys that are tuples.

Example: prevalence and testing of AIDS

	AIDS	
TEST	true	false
positive	0.003648	0.022915
negative	0.000052	0.973385
	0.003700	0.996300

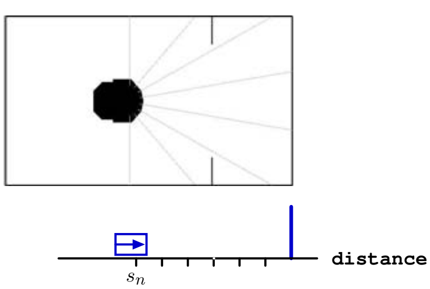
```
>>> AIDS = dist.DDist('true':0.0037, 'false':0.9963)
>>> AIDSandTEST = dist.JDist(AIDS,TESTgivenAIDS)
DDist((false,negative):   0.973385,
      (true,positive):    0.003648,
      (true, negative):   0.000052,
      (false,positive):   0.022915)
```

Applying Probability to Robot Navigation

Where am I?

– based on my current velocity

– based on noisy sensors



Hidden Markov Models

System with a state that changes over time, probabilistically.

- Discrete time steps $0, 1, \ldots, t$
- Random variables for states at each time: $S_0, S_1, S_2, \ldots$
- Random variables for observations: $O_0, O_1, O_2, \ldots$

State at time t determines the probability distribution:

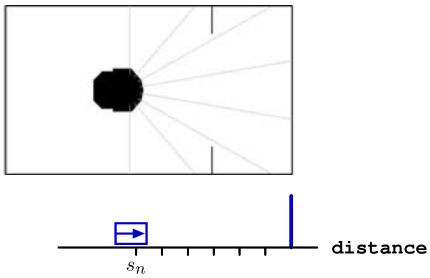
- over the observation at time t
- over the state at time $t+1$

- **Initial state distribution:**
$$\Pr(S_0 = s)$$
- **State transition model:**
$$\Pr(S_{t+1} = s \mid S_t = r)$$
- **Observation model:**
$$\Pr(O_t = o \mid S_t = s)$$

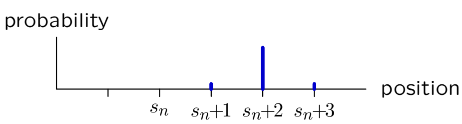
Inference problem: given sequence of observations $o_0, \ldots, o_t$, find
$$\Pr(S_{t+1} = s \mid O_0 = o_0, \ldots, O_t = o_t)$$

Transition Model

Based on my velocity and where I think I am, my next location will be ...

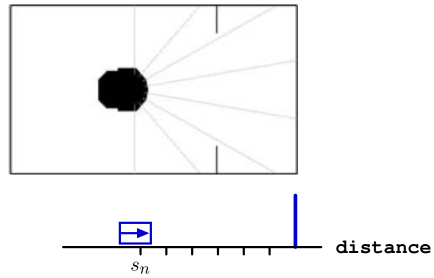


State Transition Model: probability of next state given current state

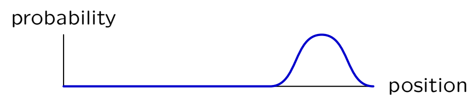


Observation Model

Based on the sonars, I am at ...



Observation Model: probability of sonar reading given current state



Hidden Markov Models

System with a state that changes over time, probabilistically.

- Discrete time steps $0, 1, \ldots, t$
- Random variables for states at each time: $S_0, S_1, S_2, \ldots$
- Random variables for observations: $O_0, O_1, O_2, \ldots$

State at time t determines the probability distribution:

- over the observation at time t
- over the state at time $t+1$

- **Initial state distribution:**

$$\Pr(S_0 = s)$$

- **State transition model:**

$$\Pr(S_{t+1} = s \mid S_t = r)$$

- **Observation model:**

$$\Pr(O_t = o \mid S_t = s)$$

Inference problem: given sequence of observations $o_0, \ldots, o_t$, find

$$\Pr(S_{t+1} = s \mid O_0 = o_0, \ldots, O_t = o_t)$$

What About the Bet?

Let's Make a Deal:

- There are four lego bricks in a bag.
- The lego bricks are either white or red.
- You get to pull one lego brick out of the bag.
- I give you $\begin{cases} \$20 & \text{if the brick is red} \\ \$0 & \text{otherwise} \end{cases}$

How much would you pay to play this game?

What About the Bet?

Which legos could be in the bag?

- 4 white
- 3 white + 1 red
- 2 white + 2 red
- 1 white + 3 red
- 4 red

How likely are these?
Assume equally likely (for lack of a better assumption)

$s = \#$ of red	0	1	2	3	4
$\Pr(S = s)$	1/5	1/5	1/5	1/5	1/5
E($\$|S = s$)	\$0.00	\$5.00	\$10.00	\$15.00	\$20.00
E($\$, S = s$)	\$0.00	\$1.00	\$2.00	\$3.00	\$4.00
E(\$)			\$10.00		

Thinking About Additional Information Quantitatively

Assume that a red lego is pulled from the bag and then returned.

How much money should you now expect to make?
We need to update the state probabilities.

$s = \#$ of red	0	1	2	3	4
$\Pr(S = s)$	1/5	1/5	1/5	1/5	1/5
$\Pr(O_0 = \text{red} \mid S = s)$	0/4	1/4	2/4	3/4	4/4
$\Pr(O_0 = \text{red}, S = s)$	0/20	1/20	2/20	3/20	4/20
$\Pr(S = s \mid O_0 = \text{red})$	0/10	1/10	2/10	3/10	4/10
E($\$|S = s$)	\$0.00	\$5.00	\$10.00	\$15.00	\$20.00
E($\$, S = s \mid O_0 = \text{red}$)	\$0.00	\$0.50	\$2.00	\$4.50	\$8.00
E($\$ \mid O_0 = \text{red}$)			\$15.00		

These are examples of precise statements about uncertain situations.

Thinking About Additional Information Quantitatively

Assume that a white lego is pulled from the bag and then returned.

How much money should you now expect to make?
We need to update the state probabilities.

$s = \#$ of red	0	1	2	3	4
$\text{Pr}(S = s)$	$1/5$	$1/5$	$1/5$	$1/5$	$1/5$
$\text{Pr}(O_0 = \text{white} \mid S = s)$	$4/4$	$3/4$	$2/4$	$1/4$	$0/4$
$\text{Pr}(O_0 = \text{white}, S = s)$	$4/20$	$3/20$	$2/20$	$1/20$	$0/20$
$\text{Pr}(S = s \mid O_0 = \text{white})$	$4/10$	$3/10$	$2/10$	$1/10$	$0/10$
E($ | $S = s$)	\$0.00	\$5.00	\$10.00	\$15.00	\$20.00
E($, $S = s \mid O_0 = \text{white}$)	\$0.00	\$1.50	\$2.00	\$1.50	\$0.00
E($ | $O_0 = \text{white}$)			\$5.00		

These are examples of precise statements about uncertain situations.