ASEN 5044, Fall 2018

Statistical Estimation for Dynamical Systems

Lecture 8:
Intro to Probability

Prof. Nisar Ahmed (Nisar.Ahmed@Colorado.edu)

Thurs 9/20/2018

University of Colorado Boulder

Aerospace Engineering

Announcements

- **HW 3 Due Thurs 9/27 at 11 am (before start of next lecture)**
- Submit to Canvas

Overview

- <u>Last Time:</u> Stability and Observability of LTI DT systems

$$\mathcal{O} = \begin{bmatrix} H \\ HF \\ \vdots \\ HF^{n-1} \end{bmatrix} \longrightarrow \text{rank } \mathcal{O} \stackrel{?}{=} n \text{ (full col rank)} \longrightarrow \text{if yes, then } (H,F) \text{ observable}$$

$$\longrightarrow \text{can use observability } \underbrace{\mathcal{O}^T\mathcal{O}}_{\text{Gramian}} \text{ to find } \underline{\underline{x(0)}} \text{ if obs.}$$

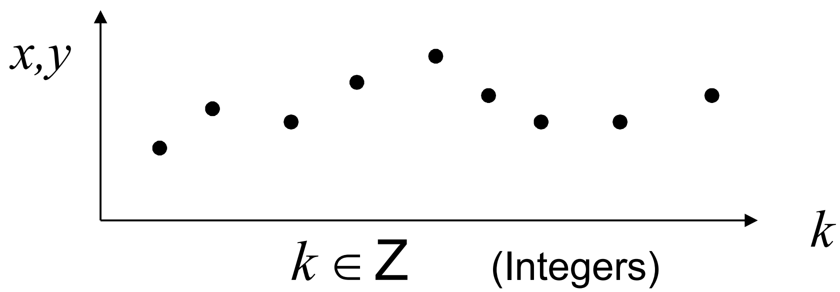
- <u>Today:</u> Intro to Probability
    - Motivation
    - Formal definitions: sample spaces, event spaces, axioms…
    - Basic operations

READ: Chapter 2.3 in Simon book

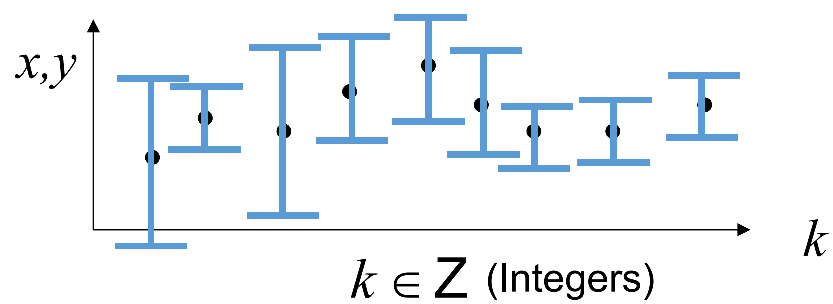
Intro to Probability

- How to describe **random non-determinism (i.e. random errors/noise)** in SS models?
- Need to use **probability theory**
  - What are probabilities?
  - What are the rules?
  - How to interpret for (scalar/vector-valued) quantities?

(Deterministic variables in DT)



(Random non-deterministic variables in DT)



Example: 1D Robot Localization with an Inertial Sensor

- Consider robot with inertial position p(t) and acceleration input u(t)
- Robot can measure position with a GPS sensor
- Suppose we take CT model, and discretize with ZOH u(t) and time step Δt

$$x = \begin{bmatrix} p \\ \dot{p} \end{bmatrix} \Rightarrow \dot{x} = \begin{bmatrix} \dot{p} \\ \ddot{p} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \quad [\text{CT ODE}]$$

$$\xrightarrow[\text{ZOH } \Delta t]{\text{DT}} \quad x(k+1) = \begin{bmatrix} p(k+1) \\ \dot{p}(k+1) \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}}_{F\ (\text{STM})} x(k) + \underbrace{\begin{bmatrix} \frac{1}{2}\Delta t^2 \\ \Delta t \end{bmatrix}}_{G} u(k)$$

$$y(k+1) = [1 \quad 0]\, x(k)$$

- But robot needs to determine its <u>actual position(k)</u> <u>**and** velocity(k)</u> to drive on road correctly!
- Can position and velocity states at any time step k both be determined if only u and y are available?

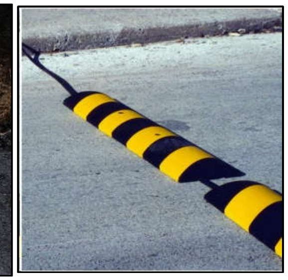
→ Check $\mathcal{O}$ [observability matrix]

$$\mathcal{O} = \begin{bmatrix} H \\ HF \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & \Delta t \end{bmatrix} \rightarrow \text{rank}(\mathcal{O}) = 2 \overset{\checkmark}{=} n \ (\text{observable})$$

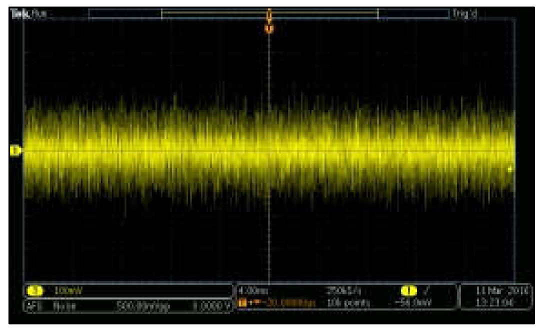
Example: 1D Robot Localization with an Inertial Sensor

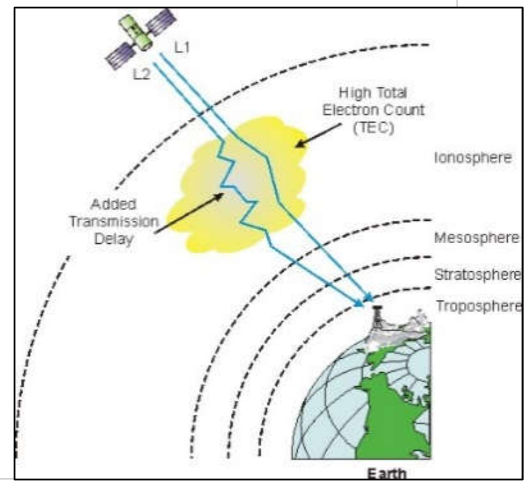
- But how to account for all of the random physical disturbances that robot encounters?

$$x(k+1) = Fx(k) + Gu(k) + w(\overset{k}{\cancel{k+1}})$$



$$y(k+1) = Hx(k+1) + Du(k+1) + v(k+1)$$





Probability as a Description of Randomness

- Probability is a **mathematical tool** for describing **random events, i.e.**
  - **those events which we do not/cannot explicitly know to have occurred for certain…**
  - **…but which would otherwise "explain" why certain other observed events happen**
- "…probability **summarizes uncertainty due to our laziness and ignorance**"
  (Stuart Russell and Peter Norvig, "Artificial Intelligence")

- Laziness: too much cost/work to predict system behavior exactly (e.g. radio environment)

- Theoretical ignorance: incomplete domain knowledge (e.g. biomedical; quantum: HUP)

- Practical ignorance: even if we knew all dynamics, may not know values for all parts
  (e.g. system ID for robot actuators, solar radiation pressure, wind speed, lateral stability derivatives of aircraft)

Probability as a Description of **Randomness**

- Many different interpretations of probability exist – each has its own issues

(i) "relative frequency": prob. of event A [assumes repeatability] $= \lim_{N \to \infty} \frac{N_A}{N} = \frac{\text{\# times A occurs}}{\text{total \# of expts.}}$

→ but in what sense does the limit exist? (ie can't actually afford to run ∞ expts.!)

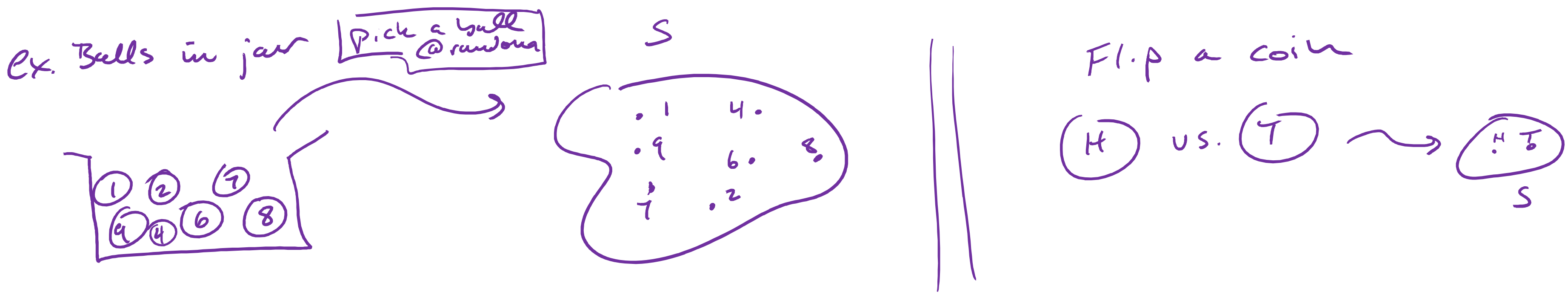
-(ii) "measure of belief" (subjective): make an argument that may or may not be based on experimental data (but they're based on good old common sense)

Formal Axiomatic Definition of Probability (modern underpinning)

- Define an **experiment** to be a **process with a random outcome**

(i.e. actual outcome is **unpredictable: we can't know what will actually happen,** but the important point is that **we do know what things could possibly happen**)

- **Sample space S**: the set of all possible random outcomes of the experiment

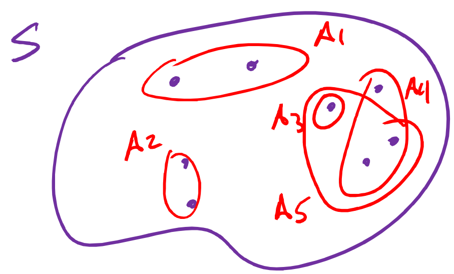


- The sample space S is also sometimes called the **outcome space**

Events and Event Spaces

- **Event**: a subset A of outcomes for the experiment, i.e. A is subset of sample space S
- **Event space**: the set of all events A that are defined on S

- Many ways to define events: each element of S could form its own event, or multiple elements of S could form events, or we can mix and match…



Event i : A_i

Event space : $\{A_1, A_2, \ldots, A_n\}$
[n events]

- If we observe an outcome in S belonging to event A, then we say that **event A has occurred** (or, **event A has been observed**)

Probabilities and Probability Spaces

- **3 Axioms of Probability**: The **probability of event A**, denoted P(A), satisfies:

(1) $P(A) \geq 0$

(2) $P(S) = 1$

(3) IF $A \cap B$ = set of outcomes in both events A & B & $A \cap B = \phi$ [empty set]

[$A \cup B$: set of outcomes in either A or B]

then $P(A \cup B) = P(A) + P(B)$

otherwise: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ [i.e $P(A \cap B) \neq 0$]

[i.e $A \cap B \neq \phi$]

- **Probability space:** a given specification $(S, \{A_i\}, \{P(A_i)\})$ (i.e. a specific sample space S, an event space on S, and probabilities on that event space)
- **This formalism covers both discrete and continuous sample spaces S** (we will focus on discrete case for time being, and move to continuous case later…)

Example: Rolling a Die

- Consider observing the number on top of a standard 6-sided die

→ set of outcomes: $\underline{S} = \{ f_1, f_2, \ldots, f_6 \}$

f_i = "outcome = face i on top" $(i = 1, \ldots, 6)$

Events: are subsets of S, e.g.

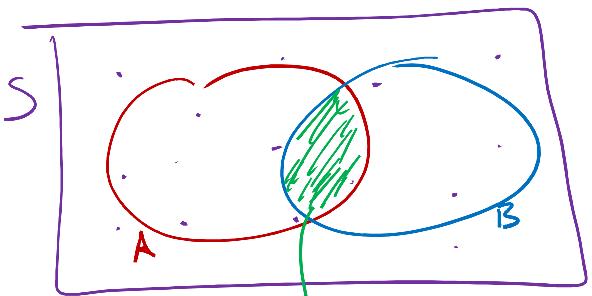
- event $F_1 = \{ f_1 \}$
- event $F_{even} = \{ i = \text{even} \} = \{ f_2, f_4, f_6 \}$
- event $F_{☺} = \{ 2 < i \leq 4 \} = \{ f_3, f_4 \}$
- sure event : $\{ \text{any face} \} = S$
- Impossible : $\{ i \leq 1 \ \& \ \text{even} \} = \emptyset \rightarrow P(\ldots) = 0$

From axioms, $p(F_{even}) = \sum_{i=2,4,6} p(f_i) = \frac{1}{2}$

⇒ probabilities are postulated for each element of S, e.g. $p(f_i) = \frac{1}{6}$ (fair die)

Joint Distributions

- The joint probability P(A,B) is the probability that events A and B occur *simultaneously*



$A \cap B = A \,\&\, B = (A,B)$

$P(A,B) = P(A \cap B) = P(A \,\&\, B)$

→ must have (from axioms):

$$\sum_a \sum_b P(A=a, B=b) = 1$$ for all possible outcomes a for A & b for B

Note (in general): $P(A=a) \cdot P(B=b) \neq P(A=a, B=b)$

[i.e. $P(A=a) \cdot P(B=b)$ may equal $P(A=a, B=b)$ in certain cases, but not always]

Example: Die Rolls

- Consider rolling of a fair 6-sided die with outcome i
- Let event A=1 if roll comes up even (i = 2,4,6) and A=0 otherwise
- Let event B=1 if roll comes up prime (i = 2,3,5) and B=0 otherwise

→ Form possible joint outcomes for $A=a$ & $B=b$, where $a \in \{0,1\}$, $b \in \{0,1\}$

A	B	occurs if ...
0	0	$i=1$ $(P_i = 1/6)$ [Not even & not prime]
0	1	$i=3,5$ $(P_i = \frac{1}{6}+\frac{1}{6} = \frac{2}{6})$ [Not even & is prime]
1	0	$i=4,6$ $(P_i = \frac{1}{6}+\frac{1}{6} = \frac{2}{6})$ [is even & not prime]
1	1	$i=2$ $(P_i = 1/6)$

Joint Probability Table (JPT)

$P(A \cap B)$

A \ B	b=0	b=1
a=0	1/6	2/6
a=1	2/6	1/6

→ Sum over rows & cols $\overset{\checkmark}{=} 1$

$\left(\text{i.e. } \sum_a \sum_b P(A=a \,\&\, B=b) \overset{\checkmark}{=} 1\right)$

Note: $P(A=1) = P(i \text{ is even}) = 1/2$

$P(B=1) = P(i \text{ is prime}) = 1/2$

$P(A=1)\cdot P(B=1) = \frac{1}{2}\cdot\frac{1}{2} = 1/4 \neq 1/6$

Marginal Probabilities

- How are P(A) and P(B) related to P(A,B)?
- Turns out we can always recover P(A) and P(B) from the joint probability P(A,B), (even though **in general** $P(A)P(B) \neq P(A,B)$)

Given joint probabilities $P(A = a, B = b)$, define:

Marginal distribution of A: $P(A = a) = \sum_b P(A = a, B = b)$

Marginal distribution of B: $P(B = b) = \sum_a P(A = a, B = b)$

Also holds for more than 2 events, e.g. if 3rd event C with $P(A, B, C)$, then

$$P(A = a) = \sum_b \sum_c P(A = a, B = b, C = c),$$

$$P(B = b) = \sum_a \sum_c P(A = a, B = b, C = c),$$

$$P(C = c) = \sum_a \sum_b P(A = a, B = b, C = c)$$

Examples of Marginal Probabilities

- 6-sided die example again: A: roll is even # (1=yes, 0 = no)

  B: roll is prime # (1=yes, 0 = no)

P(A & B)	B=0	B=1
A=0	1/6	2/6
A=1	2/6	1/6

Marginal Prob of $A = 0$: $P(A = 0) = \sum_b P(A = 0, B = b)$
$= P(A = 0, B = 0) + P(A = 0, B = 1) = \frac{1}{6} + \frac{2}{6} = \frac{1}{2}$

Marginal Prob of $A = 1$: $P(A = 1) = \sum_b P(A = 1, B = b)$
$= P(A = 1, B = 0) + P(A = 1, B = 1) = \frac{2}{6} + \frac{1}{6} = \frac{1}{2}$

Likewise, can show that:

$$P(B = 0) = \sum_a P(A = a, B = 0) = \frac{1}{2} \qquad P(B = 1) = \sum_a P(A = a, B = 1) = \frac{1}{2}$$