Probability for Estimation (review)

In general, we want to develop an estimator for systems of the form:

$$\left.\begin{aligned} \dot{x} &= f(x,u) + \eta(t); \\ y &= h(x) + \omega(t); \end{aligned}\right\} \quad given\ y, find\ \hat{x}$$

We will primarily focus on *discrete time linear systems*

$$x_{k+1} = A_k x_k + B_k u_k + \eta_k;$$

$$y_k = H_k x_k + \omega_k;$$

Where

- A_k, B_k, H_k are constant matrices
- x_k is the state at time t_k; u_k is the control at time t_k
- η_k ,ω_k are “disturbances” at time t_k

Goal: develop procedure to model disturbances for estimation

- Kolmogorov probability, based on axiomatic set theory
- 1930’s onward

Axioms of Set-Based Probability

Probability Space:

- Let Ω be a set of experimental outcomes (e.g., roll of dice)

$$\Omega = \{A_1, A_2, \dots, A_N\}$$

  - the A_i are “elementary events” and subsets of Ω are termed “events”
  - Empty set $\{\emptyset\}$ is the “impossible event”
  - $S=\{\Omega\}$ is the “certain event”

- A probability space $(\Omega, \mathrm{F}, \mathrm{P})$
  - F = set of subsets of Ω, or “events”, P assigns probabilities to events

Probability of an Event—the Key Axioms:

- Assign to each A_i a number, P(A_i), termed the “probability” of event A_i
- P(A_i) must satisfy these axioms
  1. P(A_i) ≥ 0
  2. P(S) = 1
  3. If events $A, B \in \Omega$ are “mutually exclusive,” or disjoint, elements or events ($A \cap B = \{\emptyset\}$), then

$$\mathrm{P}(A \cup B) = P(A) + P(B)$$

Axioms of Set-Based Probability

As a result of these three axioms and basic set operations (e.g., DeMorgan's laws, such as $\overline{A \cup B} = \overline{A} \cap \overline{B}$)

- P({Ø})=0
- P(A) = 1-P($\overline{A}$) $\Rightarrow$ $\mathrm{P(A)} + \mathrm{P}(\overline{A}) = 1$, where $\overline{A}$ is complement of A
- If $A_1, A_2, \dots, A_N$ mutually disjoint

$$\mathrm{P}(A_1 \cup A_1 \cup \cdots \cup A_N) = P(A_1) + P(A_1) + \cdots + P(A_N)$$

For Ω an infinite, but countable, set we add the "Axiom of infinite additivity"

3(b). If $A_1, A_2, \dots$ are mutually exclusive,

$$\mathrm{P}(A_1 \cup A_1 \cup \cdots) = P(A_1) + P(A_1) + \cdots$$

We assume that all countable sets of events satisfy Axioms 1, 2, 3, 3(b)

But we need to model uncountable sets…

Continuous Random Variables (CRVs)

Let $\Omega = \mathbb{R}$ (an uncountable set of events)

- *Problem:* it is not possible to assign probabilities to subsets of $\mathbb{R}$ which satisfy the above Axioms
- *Solution:*
  - let "events" be intervals of $\mathbb{R}$: $\mathrm{A} = \{x \mid x_l \leq x \leq x_u\}$, and their <u>*countable*</u> unions and intersections.
  - Assign probabilities to these events

$$\mathrm{P}(x_l \leq x \leq x_u) = Probability\ that\ x\ takes\ values\ in\ [x_l, x_u]$$

  - x is a *"continuous random variable* (CRV).

Some basic properties of CRVs

- If x is a CRV in $[L, U]$, then P($L \leq x \leq L$) = 1
- If y in $[L, U]$, then P($L \leq y \leq x$) = 1 - P($y \leq x \leq U$)

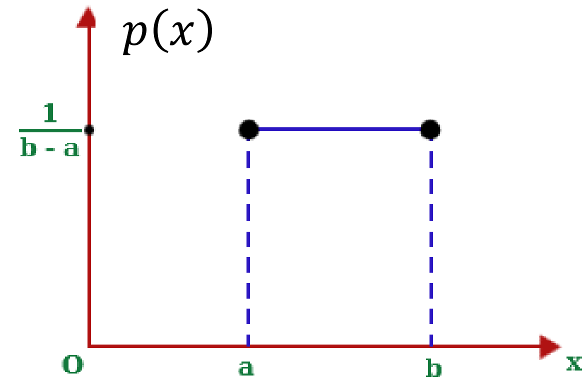
Probability Density Function (pdf)

$$p(x_l \leq x \;\leq\; x_u) \equiv \int_{x_l}^{x_u} p(x)dx$$

E.g.

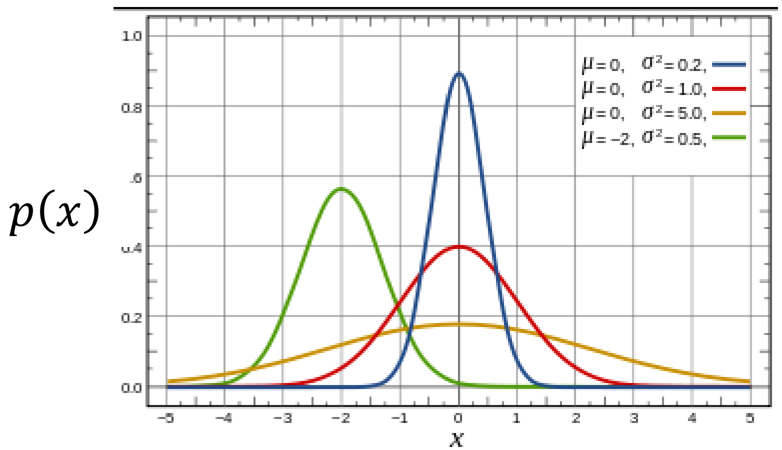
– *Uniform Probability pdf:*

$$p(x) = \frac{1}{b-a}$$



– *Gaussian (Normal) pdf:*

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$



μ =*"mean" of pdf*

σ = *"standard deviation"*

Most of our Estimation theory will be built on the Gaussian distribution

Joint & Conditional Probability

Joint Probability:

- *Countable set of events:* $\mathrm{P}(A \cap B)$ = P(A,B), probability A & B both occur
- *CRVs:* let x,y be two CRVs defined on the same probability space. Their *"joint probability density function"* $p(x,y)$ *is defined as:*

$$\mathrm{P}(x_l \leq x \leq x_u; y_l \leq y \leq y_u) \equiv \int_{y_l}^{y_u} \int_{x_l}^{x_u} p(x,y)dx\ dy$$

- *Independence*
  - A, B are independent if $\mathrm{P(A,B)} = \mathrm{P(A)\,P(B)}$
  - x,y are independent if $p(x,y) = p(x)\,p(y)$

Conditional Probability:

- *Countable events:* $\mathrm{P(A|B)} = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}$, probability of A given that B occurred
  - E.G. probability that a "2" is rolled on a fair die given that we know the roll is even:
    - $\mathrm{P(B)}$ = probability of even roll = 3/6=1/2
    - $\mathrm{P}(A \cap B) = 1/6$ (since $A \cap B = A$)
    - $\mathrm{P(2|even\ roll)} = \mathrm{P}(A \cap B)/\mathrm{P(B)} = (1/6)/(1/2) = 1/3$

Conditional Probability & Expectation

Conditional Probability (continued):

- *CRV:* $p(x|y) = \begin{cases} \frac{p(x,y)}{p(y)} & if \;\; 0 < p(y) < \infty \\ 0 & otherwise \end{cases}$
- This follows from:

$$\mathrm{P}(x_l \le x \;\le\; x_u \mid y) \equiv \int_{x_l}^{x_u} p(x|y)dx = \frac{\int_{x_l}^{x_u} p(x,y)dx}{p(y)}$$

- and:

$$p(x) = \int_{-\infty}^{\infty} p(x,y)dy = \int_{-\infty}^{\infty} p(x|y)p(y)dy$$

Expectation: (key for estimation)

- Let x be a CRV with distrubution $p(x)$. The expected value (or mean) of x is

$$E[x] = \int_{-\infty}^{\infty} xp(x)dx \qquad E[g(x)] = \int_{-\infty}^{\infty} g(x)p(x)dx$$

- Conditional mean (conditional expected value) of x given event M:

$$E[x|M] = \int_{-\infty}^{\infty} xp(x|M)dx$$

Expectation (continued)

Mean Square:

$$E[x^2] = \int_{-\infty}^{\infty} x^2 p(x) dx$$

Variance:

$$\sigma^2 = E[(x-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 p(x) dx \qquad \mu(x) = E[x]$$

Random Processes

A stochastic system whose state is characterized a time evolving CRV, x(t), t ε [0,T].

- At each t, x(t) is a CRV
- x(t) is the "state" of the random process, which can be characterized by

$$\mathrm{P}[x_l \leq x(t) \leq x_u] = \int_{-\infty}^{\infty} p(x,t)dx$$

Random Processes can also be characterized by:

- Joint probability function

Joint probability density function

$$\mathrm{P}[x_{1l} \leq x(t_1) \leq x_{1u}; x_{2l} \leq x(t_2) \leq x_{2u}] = \int_{x_{1l}}^{x_{1u}} \int_{x_{2l}}^{x_{2u}} \underbrace{p(x_1, x_2, t_1, t_2)}_{\text{Joint probability density function}} \, dx_1 dx_2$$

- Correlation Function

Correlation function

$$E[x(t_1)x(t_2)] = \int_{-\infty}^{\infty} x_1\, x_2\, p(x_1, x_2, t_1, t_2)\, dx_1 dx_2 \equiv \underbrace{\rho(t_1, t_2)}_{\text{Correlation function}}$$

- A random process x(t) is ***Stationary*** if $p(x,t+\tau)=p(x,t)$ for all τ

Vector Valued Random Processes

$\bar{X}(t) = \begin{bmatrix} X_1(t) \\ \vdots \\ X_n(t) \end{bmatrix}$ where each $X_i(t)$ is a random process

$R(t_1, t_2)$ = "Correlation Matrix" $= E[\bar{X}(t_1)\bar{X}^T(t_2)]$

$$= \begin{bmatrix} E[X_1(t_1)X_1(t_2)] & \cdots & E[X_1(t_1)X_n(t_2)] \\ \vdots & \ddots & \vdots \\ E[X_n(t_1)X_1(t_2)] & \cdots & E[X_n(t_1)X_n(t_2)] \end{bmatrix}$$

$\sum(t)$ = "Covariance Matrix" $= E[(\bar{X}(t) - \bar{\mu}(t))(\bar{X}(t) - \bar{\mu}(t))^T]$

Where $\mu(t) = E[\bar{X}(t)]$