# Machine Learning Foundations

## Chapter 6: Probability

Harish Guruprasad Ramaswamy
IIT Madras

# Outline for Chapter 6 : Probability

6.1 : Discrete Random Variables

**6.2 : Continuous Random Variables**

6.3 : Maximum Likelihood and other advanced topics

# Outline for Chapter 6 : Probability

6.1 : Discrete Random Variables

6.2 : Continuous Random Variables

1. **Random variables**
2. Expectation, Variance
3. Multiple Random Variables
4. Uniform, Exponential, Normal
5. Convergence in probability. Laws of large numbers: Markov, Chebyshev, Hoeffding, Central limit.

6.3 : Maximum Likelihood and other advanced topics

# Chapter 6.2.1 : Basics

Sample space, examples: Experiment pick a random person and measure their height.

Events as sets

Axioms of sigma-algebra

Probability Measures

Axioms of probability measures

Continuous random variables

Examples:

Amount of time you wait for a bus,

Price of a house,

Current flowing through a a piece of wire just by thermal noise

PDFs and CDFs: Definition and Properties. PDFs have units. PMFs and CDFs don't. E.g. PDF might have unit (m)^{-1}

Examples: Uniform, right angle triangular, isosceles triangular, mixture of uniform

Conditional PDFs on an event.

Example X ~ RightTrianglular(0,1) . F_{X|A} = ? Where A = {X>1/2}.

Functions of Random variables

Y=X/2, |X| or X^2 when X Is uniform [-1,1]

# Sample Space

$$(\Omega, \mathcal{F}, P)$$

# Events as Sets and Sigma Algebra

$$\mathcal{F} \subseteq \{0,1\}^{\Omega}$$

i) $\Omega \in \mathcal{F}$

ii) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$

iii) $A_1, A_2 \cdots A_n \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{n} A_i \in \mathcal{F}$

# Probability Measures

$$P : \Omega \to \mathbb{R}_+$$

i) $P(A) \geq 0$

ii) $P(\Omega) = 1$

iii) $P(A_1 \cup \cdots \cup A_n) = \sum_{i=1}^{n} P(A_i)$ for disjoint sets $A_1 \ldots A_n$.

# Examples

$$(\Omega, \mathcal{F}, P)$$

# Continuous Random Variables

$X : \Omega \to \mathbb{R}$

Domain & Range to be uncountable.

# Examples

Experiment : Waiting for bus.

$X$ : Amount of time you wait.

7:00, 7:13, 7:30, . . .

Reach the bus stop 7:10 & 7:20.

# PDF and CDF

$X \qquad x \in \mathbb{R}$

$$f_X(x) = \frac{P(X \in [x, x+dx])}{dx} \qquad \text{PDF}$$

$$F_X(x) = P(X \leq x) \qquad \text{CDF}$$

Props:

i) $f_X(x) \geq 0$

ii) $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$

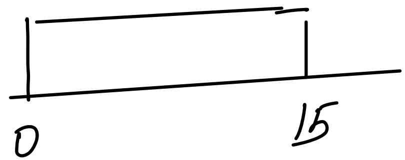
(ii) $F_X(-\infty) = 0$

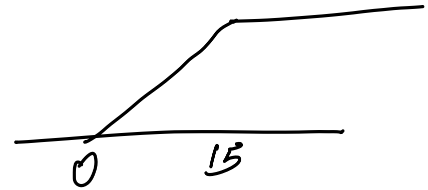
(iii) $F_X(\infty) = 1$

(iv) $F_X$ is increasing.

# Examples

Ex: $X$ is waiting time for bus



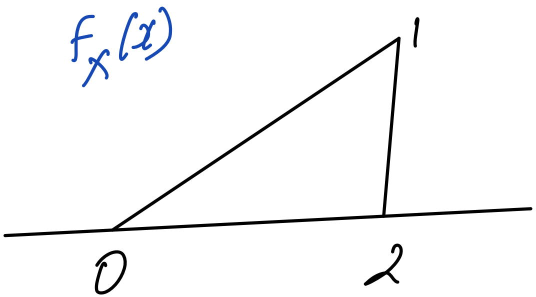



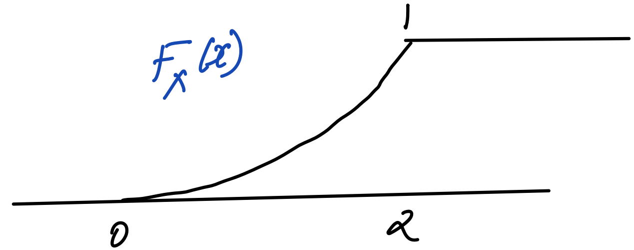

$$f_X(x) = \begin{cases} \frac{1}{15}\,(\text{min})^{-1} & \text{if } x \in [0, 15] \\ 0 & \text{otherwise.} \end{cases}$$

$$F_X(x) = \begin{cases} 0 & \text{if } x < 0 \\ x/15 & \text{if } x \in [0, 15] \\ 1 & \text{if } x > 15 \end{cases}$$

# Examples

$$f_X(x) = \begin{cases} x/2 & \text{if } x \in [0,2] \\ 0 & \text{otherwise} \end{cases}$$




$$F_X(x) = P(X \leq x)$$

$$= \int_{-\infty}^{x} P(X \in [x, x+dx])$$

$$= \int_{0}^{x} f_X(x)\,dx = \frac{1}{2}\left[\frac{x^2}{2}\right]_0^x = \frac{x^2}{4}$$