

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

1. Probability space

2. Conditioning

3. Random variables

4. Expectation and Variance

5. Multiple Random Variables

6. Bernoulli, Binomial, Poisson and Geometric RVs

6.2 : Continuous Random Variables

6.3 : Maximum Likelihood and other advanced topics

Random Variables

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

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

Example : DDTF

$$X(\omega) = \omega_1 \quad (\text{First die result})$$

$$X((a, b)) = a$$

Examples of Random Variables

Example 2: DDTE

X = Sum of the faces of dice

$$X(a,b) = a+b$$

Example 3: DLTE : $\Omega = \{HH, HT, TH, TT\}$

X = Number of heads

$$X(HH) = 2$$

$$X(HT) = 1 \quad ; \quad X(TH) = 1$$

$$X(TT) = 0$$

Probability Mass Function

$$f_X : \mathcal{R} \rightarrow [0, 1]$$

$$\begin{aligned} f_X(x) &= P(X=x) \\ &= P(\{\omega \in \Omega : X(\omega) = x\}) \end{aligned}$$

$$(i) \quad f_X(x) \in [0, 1]$$

$$(ii) \quad \sum_{x \in \text{Range}(X)} f_X(x) = 1$$

$$\begin{aligned} \sum_{x \in \text{Range}(X)} f_X(x) &= \sum_x P(\{\omega \in \Omega : X(\omega) = x\}) \\ &= P(\Omega) = 1 \end{aligned}$$

Probability Mass Function

Cumulative Distribution Function

$$F_X$$

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

Properties:

$$(i) \quad F_X(x) \in [0, 1]$$

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

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

Cumulative Distribution Function

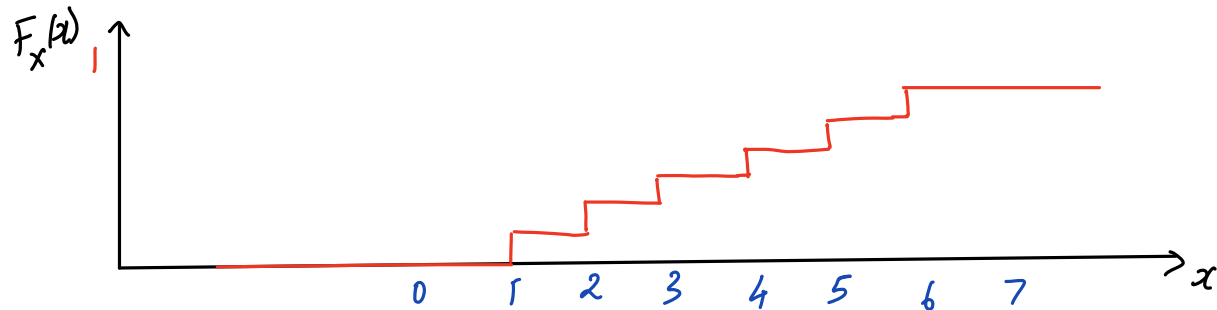
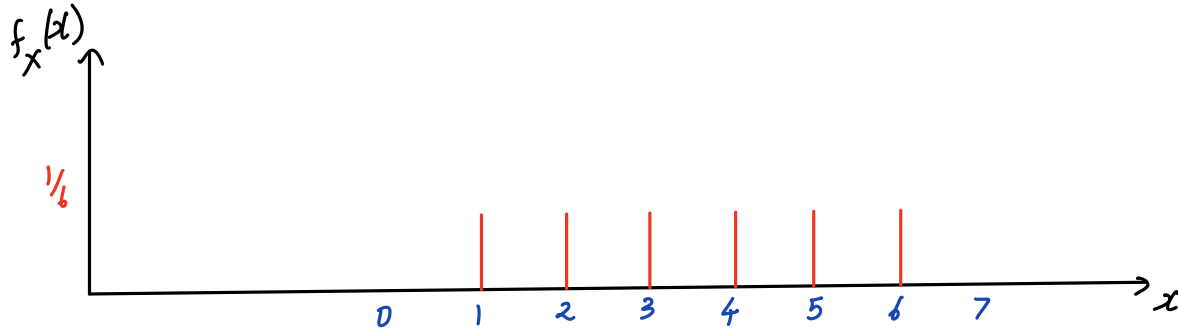
Examples of PMF and CDF

DDTE:

$$P(A) = \frac{|A|}{36}$$

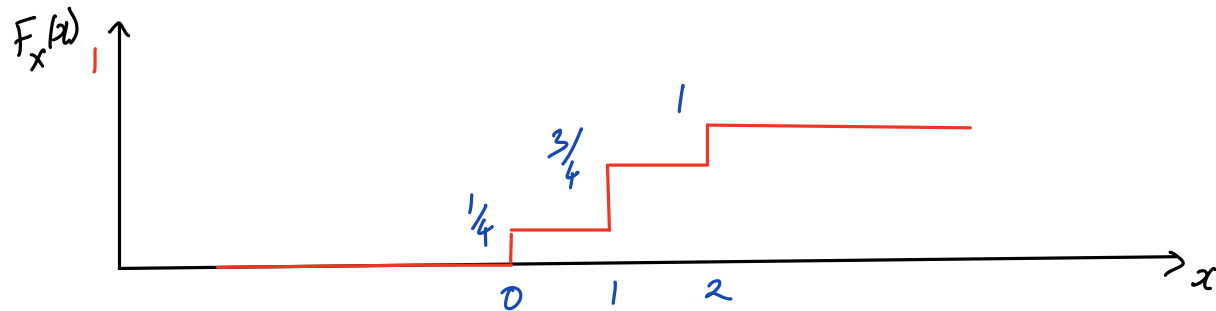
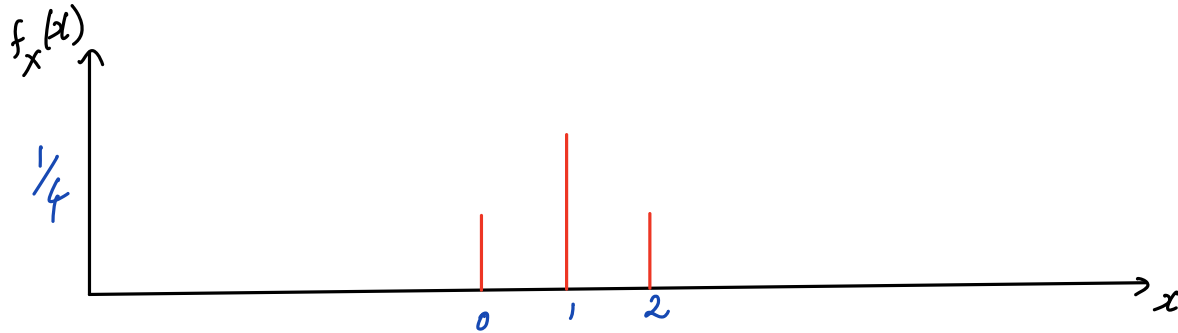
X : First face of the dice thrown

$$p(x = 0)$$



Examples of PMF and CDF

DCTE: $P(A) = \frac{|A|}{4}$. $X = \# \text{ of Heads.}$



Examples of PMF and CDF

Experiment: BDE

X = Face of die shown

$$P(X=1) = C$$

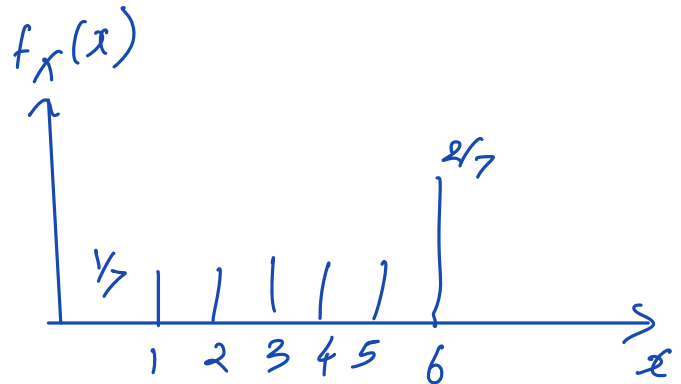
$$P(X=2) = C$$

$$P(X=3) = C$$

$$P(X=4) = C$$

$$P(X=5) = C$$

$$P(X=6) = 2C$$



$$7C = 1 \quad \Rightarrow \quad C = 1/7$$

Indicator Random Variables

$$X(\omega) = I(\omega \in A)$$

Example : DCTE

$$X = \begin{cases} 1 & \text{if 2 Heads occur} \\ 0 & \text{otherwise.} \end{cases}$$

$$P(X=0) = P(A^c) = 1 - P(A)$$

$$P(X=1) = P(A)$$

Functions of Random Variables

e.g. : DDTF

X = Result of the first die

$$Y = |X - 2|$$

$$\begin{aligned} f_Y(0) &= P(Y=0) = P(|X-2|=0) = P(X=2) \\ &= \frac{1}{6} \end{aligned}$$

$$f_Y(1) = P(X=1 \text{ or } X=3) = \frac{2}{6}$$

Examples