Foundations of Machine Learning
AI2000 and AI5000

FoML-02
Probability - Bayes Theorem and Independence

Dr. Konda Reddy Mopuri
Department of AI, IIT Hyderabad
July-Nov 2025

భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Data-driven Intelligence
& Learning Lab

So far in FoML

- What is ML?

So far in FoML

- What is ML?
- Learning Paradigms

Probability Theory

Probability Theory

- Provides a consistent framework for the quantification and manipulation of “Uncertainty”

Probability Theory

- Provides a consistent framework for the quantification and manipulation of "Uncertainty"
- Where does this 'Uncertainty' come from?

Uncertainty in ML

- Measurement Noise

Uncertainty in ML

- Measurement Noise
- Finite size of the datasets

Probability Theory

- Frequentist Interpretation

Probability Theory

- Frequentist Interpretation
  - Fraction of times the event occurs

Probability Theory

- Bayesian Approach

Probability Theory

- Bayesian Approach
  - Quantification of plausibility or strength of the belief of an event

Probability Theory

- Bayesian Approach
  - Quantification of plausibility or strength of the belief of an event
  - Modeling based approach

Probability Theory

- Bayesian Approach
  - Quantification of plausibility or strength of the belief of an event
  - Modeling based approach
  - Plays a central role in this course

Random Variable

- Stochastic variable sampled from a set of possible outcomes

Random Variable

- Stochastic variable sampled from a set of possible outcomes
- Discrete or Continuous

Random Variable

- Stochastic variable sampled from a set of possible outcomes
- Discrete or Continuous
- Probability distribution $p(X)$

Random Variable - Example (discrete)

- Throwing a dice

Random Variable - Example (discrete)

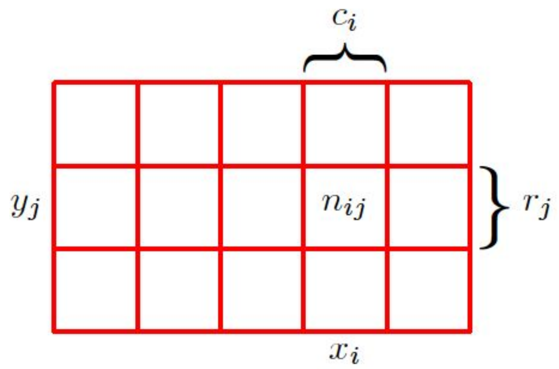
- Flipping a coin

Two Discrete Random Variables

- X
- Y

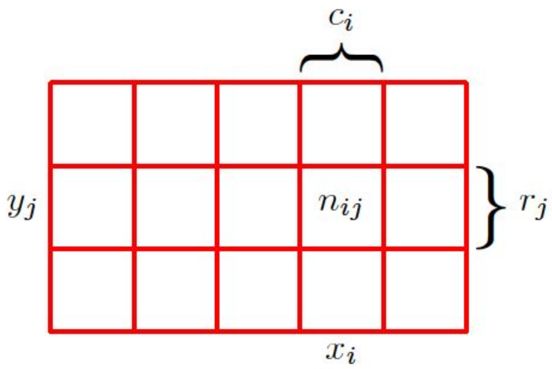
Two Discrete Random Variables

- X
- Y
- N trails: sample both



Two Discrete Random Variables

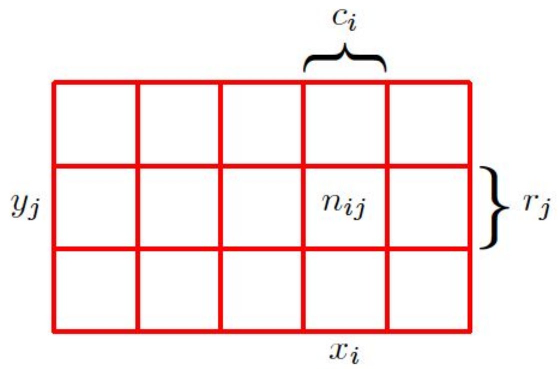
- Joint probability



Two Discrete Random Variables

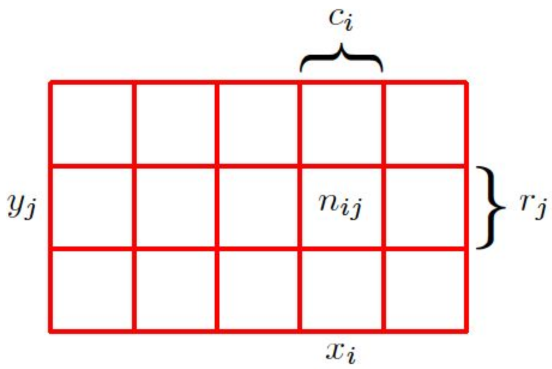
- Joint probability

$$p(X = x_i, Y = y_j) =$$



Two Discrete Random Variables

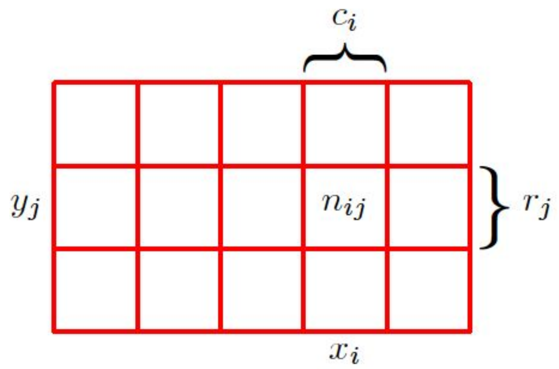
- If I am interested only on X



Two Discrete Random Variables

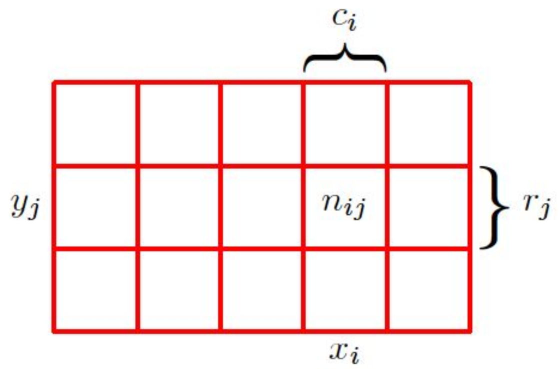
- If I am interested only on X
- Marginal probability of X

$$p(X = x_i)$$



Sum rule of Probability

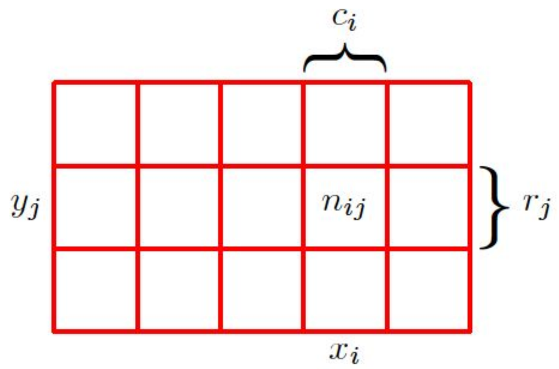
$$p(X = x_i) = \sum_{j=1}^{3} p(X = x_i, Y = y_j)$$



Conditional Probability

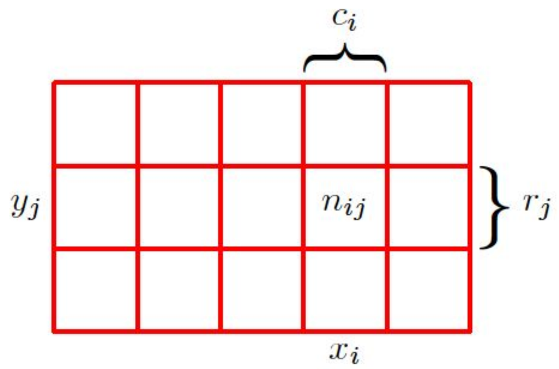
- Conditional probability of Y given X

$$p(Y = y_j / X = x_i) =$$



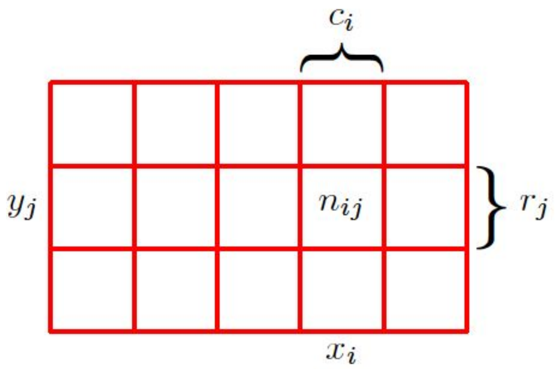
Product Rule of probability

$$p(Y = y_j / X = x_i) =$$



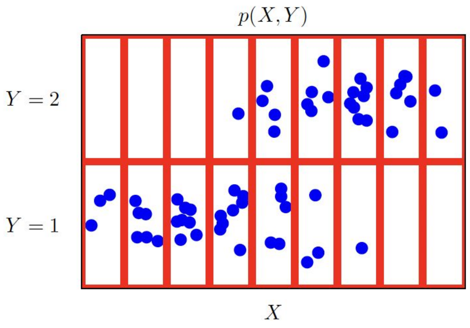
Product Rule of probability

$$p(Y = y_j / X = x_i) =$$

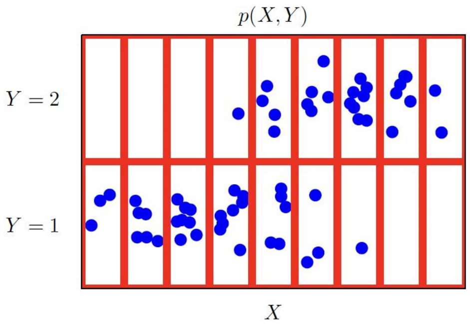


$$p(Y = y_j, X = x_i) = p(Y = y_j / X = x_i) \cdot p(X = x_i)$$

Example: Marginal & Conditional distributions

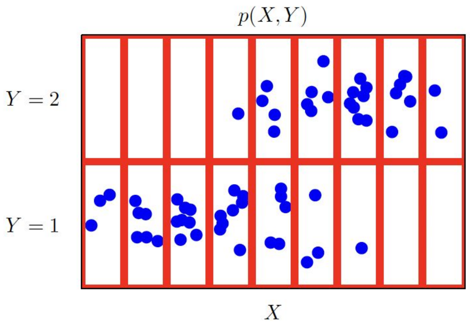


Example: Marginal & Conditional distributions



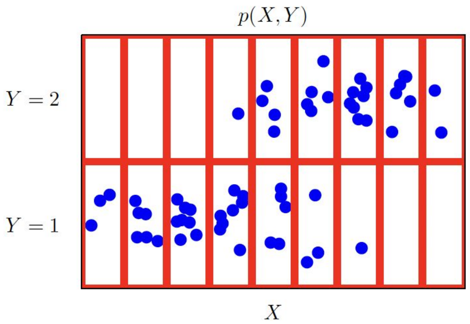
- X
- Y
- 60 trails - sample both

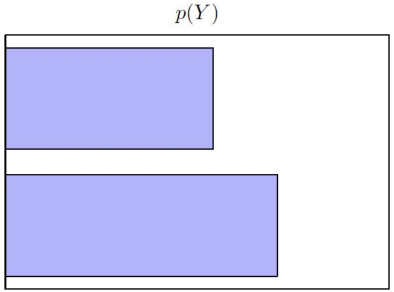
Example: Marginal & Conditional distributions



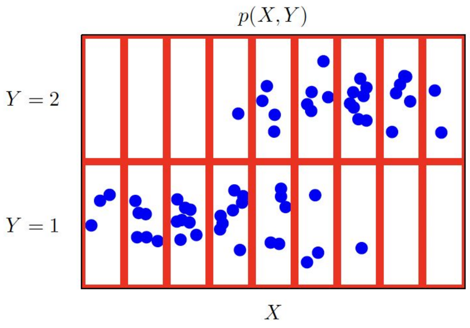
- Marginal distribution p(Y)

Example: Marginal & Conditional distributions



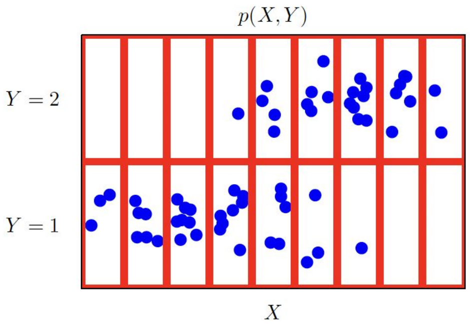


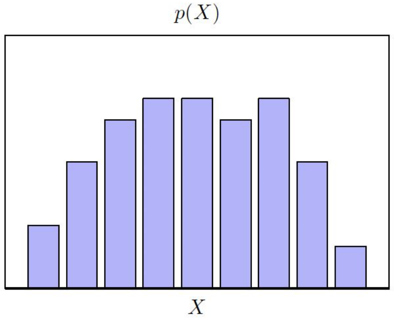
Example: Marginal & Conditional distributions



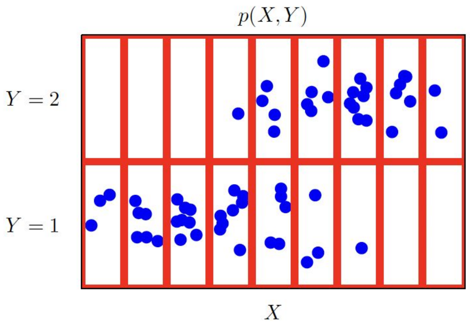
- Marginal distribution p(X)

Example: Marginal & Conditional distributions



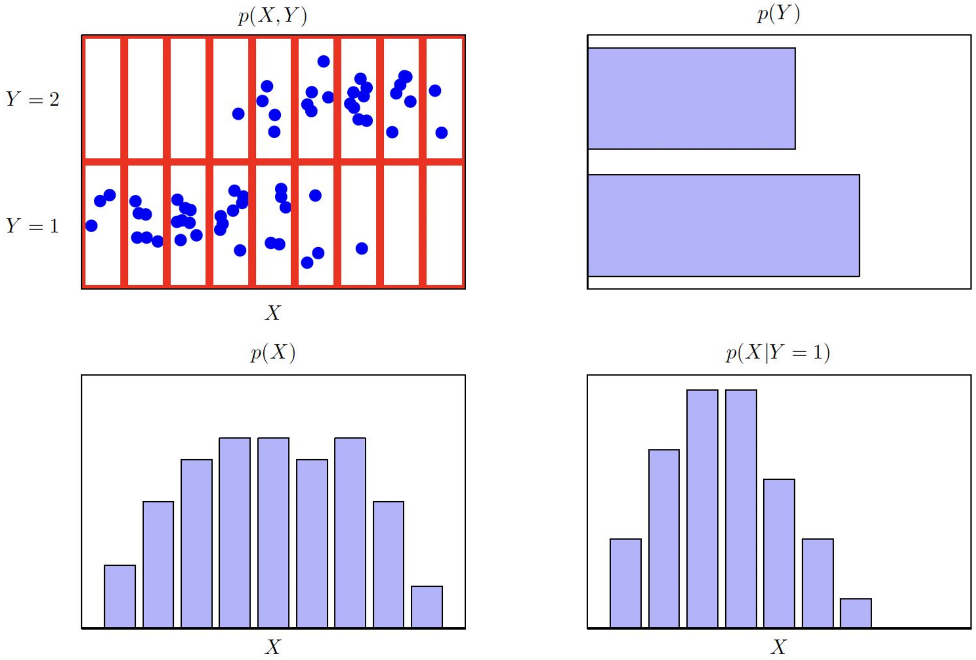


Example: Marginal & Conditional distributions



- Conditional distribution of X

Example: Marginal & Conditional distributions



Example: Marginal & Conditional distributions

$$\sum_{y \in Y} p(Y = y_i / X = x_i) = ?$$

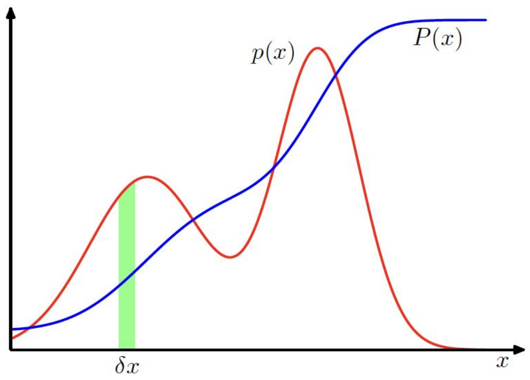
Continuous Random Variable

Continuous Random Variable

- p(X): Probability density over X

Continuous Random Variable

- p(X): Probability density over X
- Probability of x falling in (x, x+dx)
- Probability over a finite interval (a, b)



Continuous Random Variable

- Non-negativity
- Normalization

Continuous Random Variable

- Change of variables

Continuous Random Variable

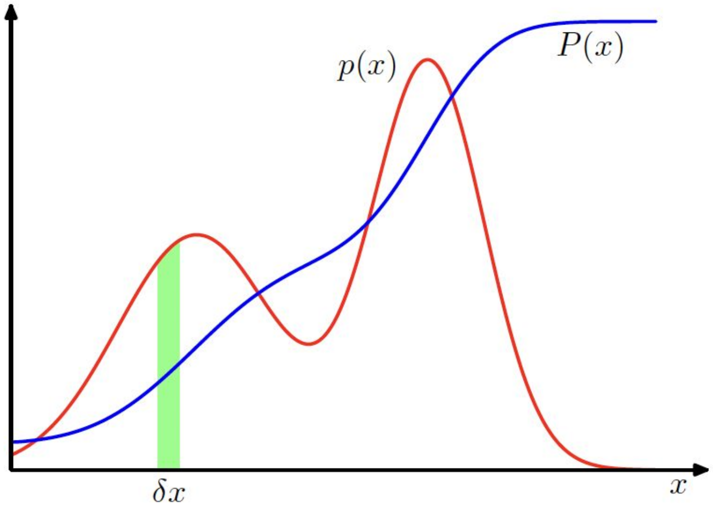
- Change of variables
- $x = g(y)$

Continuous Random Variable

- Change of variables
- $x = g(y)$
- Probabilities in $(x, x+dx)$ must be transformed to $(y, y+dy)$

Continuous Random Variable

- Cumulative distribution function



Rules of Probability Theory

	Discrete	Continuous
Additivity	$p(X \in A) = \sum_{x \in A} p(x)$	
Positivity	$p(x) \geq 0$	$p(x) \geq 0$
Normalization		$\int_{\mathcal{X}} p(x)dx = 1$
Sum Rule	$p(x) = \sum_{y \in Y} p(x,y)$	
Product Rule	$p(x,y) = p(x/y) \cdot p(y)$	$p(x,y) = p(x/y) \cdot p(y)$

Bayes theorem

భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Data-driven Intelligence
& Learning Lab

Bayes Theorem

- Product rule $$p(x, y) = p(x/y) \cdot p(y)$$

Bayes Theorem

- Product rule
- Symmetry property
- Bayes rule
- Denominator

$$p(x, y) = p(x/y) \cdot p(y)$$

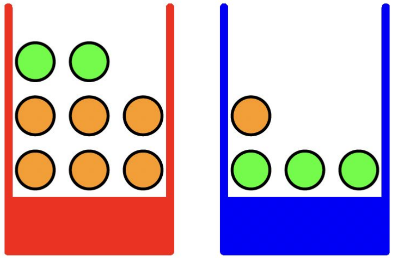
Bayes Theorem

$$p(y/x) = \frac{p(x/y) \cdot p(y)}{p(x)}$$

- Prior probability
- Posterior probability of Y
- Likelihood of X = x given Y = y
- Evidence for X = x

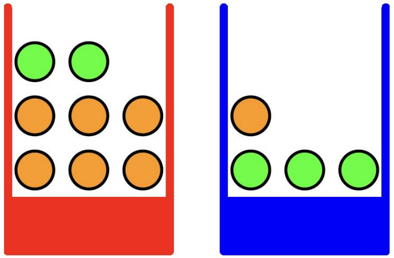
Example

Boxes and Fruits



- Random variables

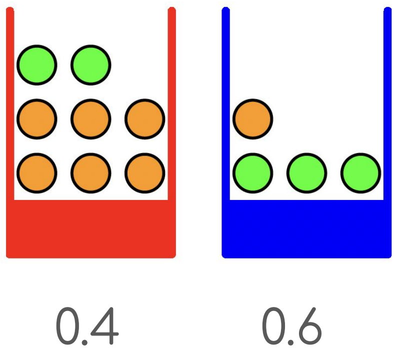
Boxes and Fruits



- Random variables
  - Box - B
  - Fruit - F

Boxes and Fruits

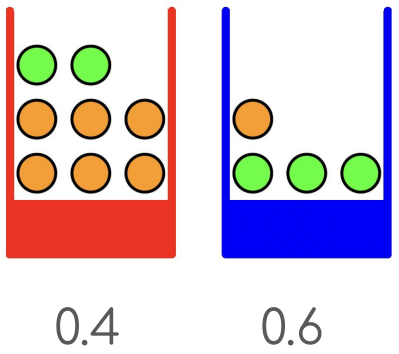
- Prior Box distribution



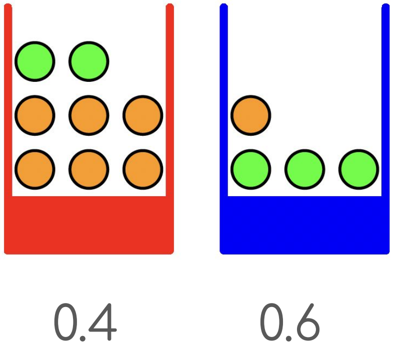
Boxes and Fruits

- Prior Box distribution
    -
- Conditional of F given B

- Marginal of F



Boxes and Fruits



- Marginals $p(F{=}a) = 11/20$ & $p(F{=}0) = 9/20$
- Posterior probability of Box given observed fruit

$$p(B = r / F = o) =$$

Independence

Independent Random variable

- Two random variables X and Y are independent iff measuring X gives no information about Y (and vice versa)

Next
Expectation, Variance, and Gaussian Distribution