Probability Theory - I

EE698V - Machine Learning for Signal Processing

Vipul Arora

Discrete Random Variables

Reference: PRML Section 1.2

Rose

The **rose** is a type of flowering shrub. Its name comes from the Latin word *Rosa*.[1] The flowers of the rose grow in many different colors, from the well-known red rose or yellow roses and sometimes white or purple roses. Roses belong to the family of plants called Rosaceae. All roses were originally wild and they come from several parts of the world, North America, Europe, northwest Africa and many parts of Asia and Oceania. There are over 100 different species of roses. The wild rose species can be grown in gardens, but most garden roses are cultivars, which have been chosen by people.[2]

Over hundreds of years they have been specially bred to produce a wide variety of growing habits and a broad range of colours from dark red to white including as well yellow and a bluish/lilac colour. Many roses have a strong, pleasant scent. Most roses have prickles(incorrectly called thorns) on their stems. Rose bushes are able to tolerate a wide variety of growing conditions. The fruit of the rose is called a hip. Some roses have decorative hips.

Source: Wikipedia

Jasmine

Jasmine is a genus of plants. They are shrubs or vines that grow in moderately warm climates. There are about 200 different species of Jasmine. They are also quite liked in gardens. Tea can be made from the flowers. Some species are used to make special oil, perfumes or incense. Women, especially from Asia sometimes wear jasmine flowers in their hair. Jasmine flowers are white or yellow in colour, although in rare instances they can be slightly reddish. These flowers are well known for their lovely smell. The Jasmine is believed to have originated in the Himalayas in western China.

Source: Wikipedia

Bag of words (selected)

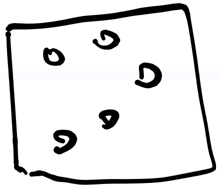
Rose	6
flower	2
color	3
Red	4
Yellow	1
White	1
Asia	1
Scent	1
Thorns	1
Latin	1

Jasmine	5
flower	3
color	1
Red	1
Yellow	1
White	4
Asia	1
Scent	1
Himalayas	1
Tea	1

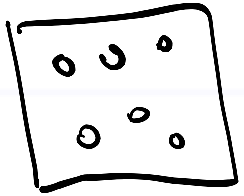
Words and their counts in each topic

Assume that
- the topic “Rose” is chosen with a probability 0.7
- the topic “Jasmine” is chosen with a probability 0.3

Examples



Rose



Jasmine

- What is probability of choosing the word “red”?
- If the chosen word is “white”, which topic did it come from?
- Does the probability of the word “asia” depend on which topic was chosen?
- Does the probability of any word depend upon which topic was chosen?

rose	6
flower	2
color	3
red	4
yellow	1
white	1
asia	1
scent	1
thorns	1
latin	1

jasmine	5
flower	3
color	1
red	1
yellow	1
white	4
asia	1
scent	1
himalayas	1
tea	1

Random Variables

- Topic: $t \in$ {Rose, Jasmine}
- Word: $w \in$ {Himalayas, Jasmine, Thorns, color, Red, Rose, Yellow, Scent, flower, Tea, Latin, Asia, White}

Joint Probability

- P(t,w)
- Example:
  - P(t=Rose, w=red)
  - P(t=Jasmine, w=asia)
- $\boxed{P(w) = \sum_t P(t,w)}$ a.k.a. Marginalization or Sum Rule

Conditional Probability

- Probability of getting a word w, given that the topic is t
- Example:
  - P(w=red | t=Rose)
- $$\boxed{P(w|t) = \frac{P(w,t)}{P(t)}}$$

Joint and conditional probabilities

- $P(w|t) = \frac{P(w,t)}{P(t)}$
- $P(w,t) = P(w|t)P(t)$
- a.k.a. Product Rule

Bayes Theorem

- $P(t|w) = \frac{P(t,w)}{P(w)} = \frac{P(w|t)P(t)}{P(w)}$
- i.e., $\boxed{P(t|w) = \frac{P(w|t)P(t)}{P(w)}}$ a.k.a. Bayes Theorem

w = Red

Back to example

Word Counts

Topic Probabilities

- P(t=Rose) = 0.7
- P(t=Jasmine) = 0.3

	Rose	Jasmine
rose	6	0
flower	2	3
color	3	1
red	4	1
yellow	1	1
white	1	4
asia	1	1
scent	1	1
thorns	1	0
latin	1	0
jasmine	0	5
himalayas	0	1
tea	0	1
	21	**19**

Example

- What is probability of choosing the word “red”?

	Rose	Jasmine
rose	6	0
flower	2	3
color	3	1
red	4	1
yellow	1	1
white	1	4
asia	1	1
scent	1	1
thorns	1	0
latin	1	0
jasmine	0	5
himalayas	0	1
tea	0	1
	21	**19**

- $\frac{4}{21} + \frac{1}{19}$
- ✓ $\frac{4}{21} \times 0.7 + \frac{1}{19} \times 0.3$
- $\frac{4+1}{40}$

$P(w = red)$

$P(t = Rose) = 0.7$; $P(t = Jas) = 0.3$

$P(w = Red \mid t = Rose) = \frac{4}{21}$

$P(w = Red \mid t = Jas) = \frac{1}{19}$

$$P(w = red) = P(w = red \mid t = rose)\, P(t = rose) + P(w = red \mid t = Jas)\, P(t = Jas)$$

$$P(w) = \sum_t P(w, t) = \sum_t P(w \mid t)\, P(t)$$

Example

- If the chosen word is "white", which topic did it come from?

$P(t = \text{Rose} \mid w = \text{white})$ & $P(t = \text{Jas} \mid w = \text{white})$

9 know: $P(t)$, $P(w \mid t)$

$$P(t \mid w) = \frac{P(t, w)}{P(w)} = \frac{P(w \mid t)\, P(t)}{\sum_t P(w \mid t)\, P(t)}$$

	Rose	Jasmine
rose	6	0
flower	2	3
color	3	1
red	4	1
yellow	1	1
white	1	4
asia	1	1
scent	1	1
thorns	1	0
latin	1	0
jasmine	0	5
himalayas	0	1
tea	0	1
	21	**19**

Example

- Does the probability of the word “asia” depend on which topic was chosen?
- Does the probability of any word depend upon which topic was chosen?

$$P(w = \text{asia} \mid t = \text{Rose}) \stackrel{?}{=} P(w = \text{asia} \mid t = \text{Jas})$$

$$\frac{1}{21} \neq \frac{1}{19}$$

Independence

if $P(w \mid t = rose) = P(w \mid t = Jasmine) = P(w)$

$P(w) = \sum_t P(w \mid t)\, P(t)$

$= P(w \mid t = rose)\, P(t = rose) + P(w \mid t = Jas)\, P(t = Jas)$

$= a$ $= a$

$= P(w) \left(\underbrace{P(t = rose) + P(t = Jas)}_{=1} \right)$

$= a$

- $\boxed{P(w|t) \neq \frac{P(w,t)}{P(t)} = P(w)}$
- $P(w, t) = P(w)P(t)$
- If both topics have similar distributions of words, then we would have independence, i.e. probability of a word does not depend on which topic was chosen

Example

- Does the probability of the word "asia" depend on which topic was chosen?
- P(w=asia|t=Rose) = $\frac{1}{21}$
- P(w=asia|t=Jasmine) = $\frac{1}{19}$
- P(w=asia) = $\frac{1}{21} \times 0.7 + \frac{1}{19} \times 0.3$

	Rose	Jasmine
rose	6	0
flower	2	3
color	3	1
red	4	1
yellow	1	1
white	1	4
asia	1	1
scent	1	1
thorns	1	0
latin	1	0
jasmine	0	5
himalayas	0	1
tea	0	1
	21	**19**

Example

- Are the random variables w and t independent?
- $P(w|t) = P(w)$?

	Rose	Jasmine
rose	6	0
flower	2	3
color	3	1
red	4	1
yellow	1	1
white	1	4
asia	1	1
scent	1	1
thorns	1	0
latin	1	0
jasmine	0	5
himalayas	0	1
tea	0	1
	21	**19**

Continuous Random Variables

Reference: PRML Section 1.2

Example

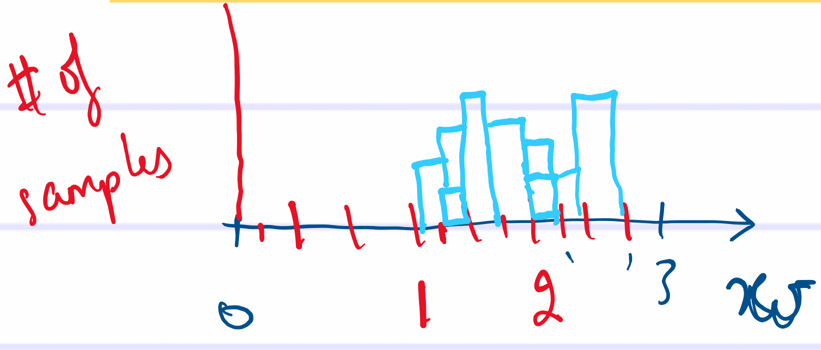
w, h

$$p(w<2 \mid h \geq 1.5) = \frac{\# \, w<2, \, h \geq 1.5}{\# \, h \geq 1.5} = \frac{1}{4}$$

- Rose size
- What is the probability to find a rose of width less than 2cm?
- What is the probability to find a rose of width less than 2cm amongst all roses of height greater than or equal to 1.50cm?
- Are width and height independent?

Sample no.	Width (cm)	Height (cm)
1	2.1	1.53
2	1.9	1.41
3	2.3	1.60
4	1.8	1.37
5	1.6	1.11
6	2.4	1.68
7	2.1	1.42
8	1.8	1.50

Estimating Joint Probabilities

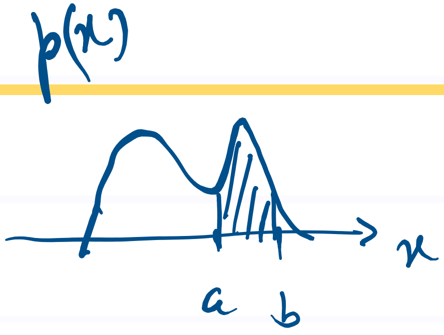


histogram

$$p\big(w \in (0.0, 2.0)\big)$$

$$p\big(w \in (2.25, 2.75)\big) = \frac{5}{\text{total \# samples}}$$

Probability Density



$\int p(x)\, dx = 1$

- $p\big(x \in (a,b)\big) = \int_a^b p(x)dx$
- Probability of finding x in a range (x, x+dx) is given by p(x)dx

Probability Distribution or CDF

cumulative distribut^n fn.

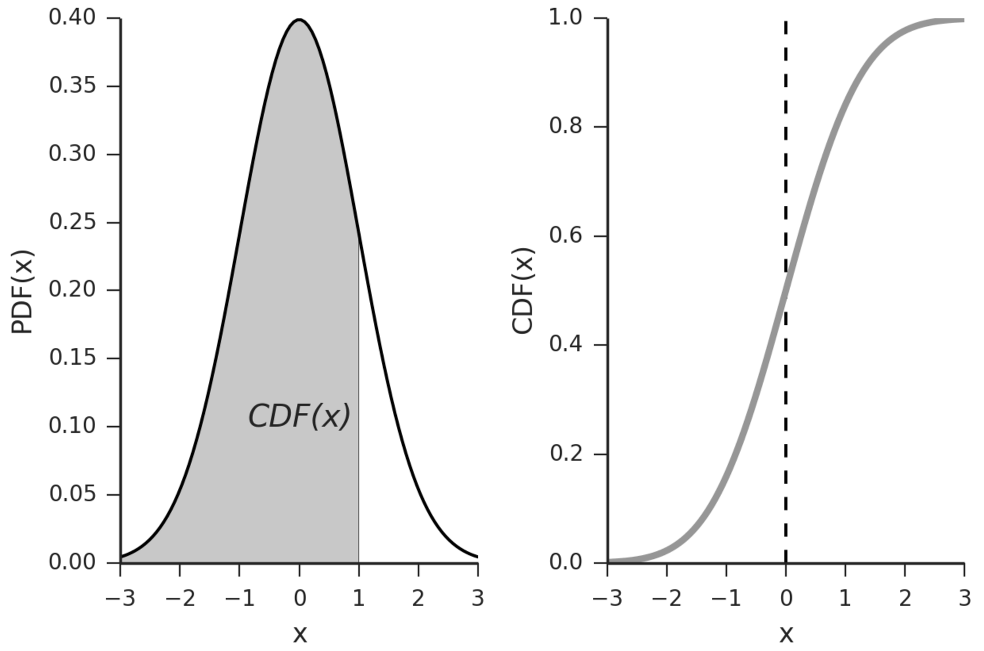
- $P(z) = \int_{-inf}^{z} p(x)dx$

$P(\infty) = 1$

$P(-\infty) = 0$

$P(z) > 0 \quad \forall z$

$\frac{dP(z)}{d(z)} = p(z) \geq 0$



Source: deepai.org

Marginalization, Conditional prob., etc.

- $p(x) = \int p(x, y)dy$ Sum Rule
- $p(x, y) = p(x|y)p(y)$ Product Rule
- $p(y|x) = \frac{p(x|y)p(y)}{p(x)}$ Bayes Rule

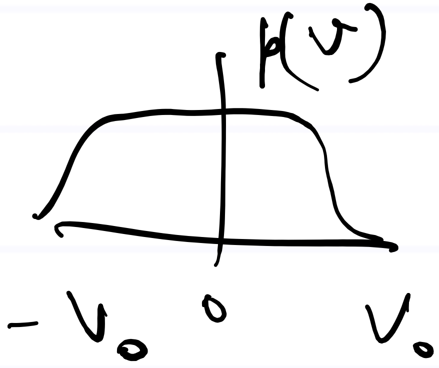
Basic properties

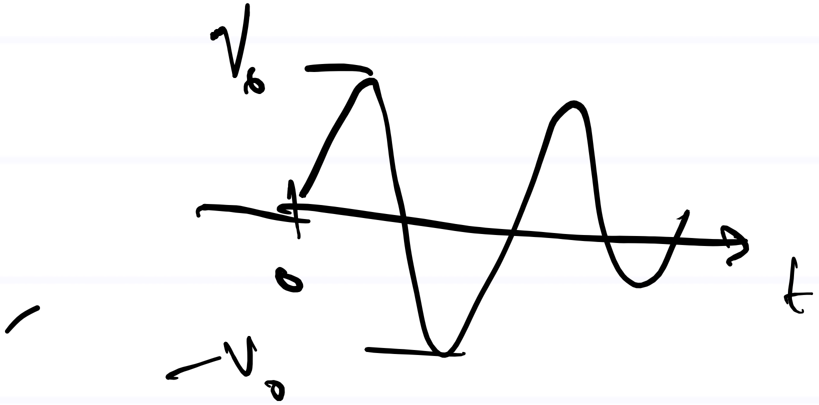
- $\int p(x)dx = 1$
- $p(x) \geq 0 \; \forall x$

✗ $p(x) \leq 1 \quad \forall x$

Transformation of Variables

- Given $p(x)$, what is $p(x^2)$?
- For voltage readings across a resistor, p(voltage) is known, what is p(power)?





$$\text{power} = \frac{v^2}{R}$$