BT-3102

COMPUTATIONAL METHODS FOR BUSINESS ANALYTICS

L1 – Revision of Probability Concepts

Aditya Karanam

Optional Readings

- Chapters 12.1 and 12.1.1: Quantifying Uncertainty

Outline

- Probability
- Random Variables
- Joint and Conditional Distributions
- Marginalization and Normalization
- Inference by Enumeration

What is this?

Uncertainty

- Let action A_t = leave for airport t minutes before flight
  - Will A_t get me there on time?

- Problems:
  - Partial observability (road state, other drivers' plans, etc.)
  - Noisy sensors (traffic reports)
  - Uncertainty in action outcomes (flat tire, etc.)
  - Immense complexity of modeling traffic (SMRT breakdowns)

Logical Approach

- Logic: Deterministic in nature (Traditional AI)
  - True or False

- Risks falsehood: "A_{15} will get me there on time"

- Conclusions are too weak for decision making:
  - "A_{15} will get me there on time *if there's no accident on the bridge*, and *if it doesn't rain*, and *if my tires remain intact*, etc."
  - A_{1440} might reasonably be said to get me there on time but I'd have to stay overnight in the airport

- Cannot enumerate all possible qualifications
  - Lazy or ignorant (don't know that we don't know)

Probabilistic Approach

- Most often we have a degree of belief
  - Running nose and coughing: What disease do I have?
- Probability is the language helps us in representing our beliefs or summarizing the uncertainty.
  - Helpful given the massive amounts of data available everywhere
  - Probability and statistics help us summarize and understand it
- Given the available evidence, A_{15} will get me there on time with probability 0.04
  - $P(A_{15} \mid \text{no reported accidents}) = 0.04$
- Probabilities change with new evidence:
  - $P(A_{15} \mid \text{no reported accidents, 5 a.m.}) = 0.15$
  - $P(A_{15} \mid \text{no reported accidents, 5 a.m., raining}) = 0.08$
    - i.e., observing evidence causes *beliefs to be updated*

Probabilistic Approach

- Not just for games of chance!
  - I'm sniffling: am I sick?
  - Tooth hurts: have cavity?
  - Email contains "FREE!": is it spam?
  - GST has increased! Will consumption grow?
  - Adding more features! Will consumers like my product?
  - Detect an object: Is it a human?

Random Variables

- An aspect of the world about which we have uncertainty
  - R = Is it raining?
  - T = Is it warm or cold?
  - D = How long will it take to drive to work?
  - L = Where am I?
- Notation: Denote random variables with capital letters
- Random variables have a domain
  - R in {true, false} or $\{r, \neg r\}$ or $\{+r, -r\}$
  - T in {warm, cold}
  - D in $[0, \infty]$
  - L in $\{(1.2966^\circ\ \text{N},\ 103.7764^\circ\ \text{E}), (\ldots, \ldots), \ldots\}$

Probability Distribution

- Associate a probability with each value of the random variable (distribution of our degree of belief over an aspect of the world)

Sunny

P(S)

S	P
rain	0.4
sun	0.6

- The probability of a lowercase value is a single number
  - P(S=rain) = 0.4
    - Shorthand notation: P(rain) same as P(S=rain)

Summing over all the values in the domain of X

- Must satisfy: (i) $\forall x \;\; P(X = x) \geq 0$ (ii) $\sum_x P(X = x) = 1$

Joint Probability Distributions

- Distribution of our degree of belief over multiple aspects of the world
- **Joint probability distribution is** over a set of random variables: $X_1, X_2, \ldots X_n$
  - A map from assignments (or *outcomes*) of random variables to real numbers

Joint Probability of Temperature and Sunny: P(T,S):

T	S	P
warm	sun	0.4
warm	rain	0.1
cold	sun	0.2
cold	rain	0.3

- Represented as: $P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$
  - Shorthand notation: $P(x_1, x_2, \ldots, x_n)$

Joint Probability Distributions

Joint Probability of T and S: P(T,S)

T	S	P
warm	sun	0.4
warm	rain	0.1
cold	sun	0.2
cold	rain	0.3

- Must obey: $0 \leq P(x_1, x_2, \ldots x_n) \leq 1$

$$\sum_{(x_1, x_2, \ldots x_n)} P(x_1, x_2, \ldots x_n) = 1$$

- Least size of distribution if n variables each having domain size of d?

$$d^n - 1$$

Event

- An *event* is a set E with assignments to random variables

$$P(E) = \sum_{(x_1 \ldots x_n) \in E} P(x_1 \ldots x_n)$$

Summing over all the values of random variables consistent with the event (E)

- From a joint distribution, we can calculate the probability of any event
- Practice (Probability of Events):
  - P(warm, sun)?
  - P(warm)?
  - P(warm or sun)?

T	S	P
warm	sun	0.4
warm	rain	0.1
cold	sun	0.2
cold	rain	0.3

Event

- An *event* is a set E with assignments to random variables

$$P(E) = \sum_{(x_1 \ldots x_n) \in E} P(x_1 \ldots x_n)$$

Summing over all the values of random variables consistent with the event (E)

- From a joint distribution, we can calculate the probability of any event

- Practice (Probability of Events):
  - P(warm, sun) 0.4
  - P(warm)? 0.4+0.1
  - P(warm or sun)? 0.4+0.1+0.2

T	S	P
warm	sun	0.4
warm	rain	0.1
cold	sun	0.2
cold	rain	0.3

Marginalization

- Marginalization (or summing out) is the process of obtaining sub-distributions (also called marginal distribution) by eliminating random variables

P(T, S)

T	S	P
warm	sun	0.4
warm	rain	0.1
cold	sun	0.2
cold	rain	0.3

$$P(t) = \sum_s P(t, s)$$

$P(T)$

T	P
warm	0.5
cold	0.5

$$P(s) = \sum_t P(t, s)$$

$P(S)$

S	P
sun	0.6
rain	0.4

$$P(X_1 = x_1) = \sum_{x_2} P(X_1 = x_1, X_2 = x_2)$$

Quiz: Marginalization

P(A, B)

A	B	P
+a	-b	0.5
+a	+b	0.1
-a	-b	0.2
-a	+b	0.2

$$P(a) = \sum_b P(a,b)$$

$$P(b) = \sum_a P(a,b)$$

P(A)

A	P
+a	
-a	

P(B)

B	P
+b	
-b	

Quiz: Marginalization

P(A, B)

A	B	P
+a	-b	0.5
+a	+b	0.1
-a	-b	0.2
-a	+b	0.2

$$P(a) = \sum_b P(a,b)$$

$$P(b) = \sum_a P(a,b)$$

P(A)

A	P
+a	0.6
-a	0.4

P(B)

B	P
+b	0.3
-b	0.7

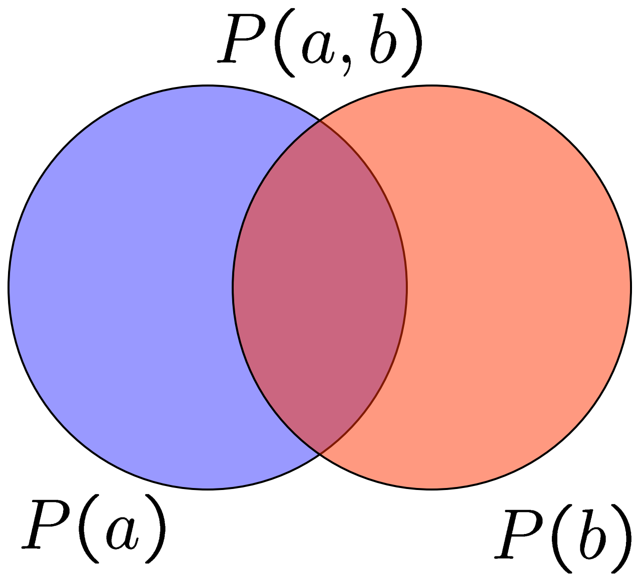
Conditional Probabilities

- A conditional probability is the probability of an event having known or observed another event (usually called as evidence)

$$P(a|b) = \frac{P(a,b)}{P(b)}$$

Event of Interest or Query

Known Event or Evidence



Practice: Conditional Probabilities

- A conditional probability is the probability of an event having known or observed another event (usually called as evidence)

$$P(a|b) = \frac{P(a,b)}{P(b)}$$

$P(T, S)$

T	S	P
warm	sun	0.4
warm	rain	0.1
cold	sun	0.2
cold	rain	0.3

$$P(S = r \mid T = c) = ?$$

Practice: Conditional Probabilities

- A conditional probability is the probability of an event having known or observed another event (usually called as evidence)

$$P(a|b) = \frac{P(a, b)}{P(b)}$$

$P(T, S)$

T	S	P
warm	sun	0.4
warm	rain	0.1
cold	sun	0.2
cold	rain	0.3

$$P(S = r \mid T = c) = \frac{P(S = r, T = c)}{P(T = c)} = \frac{3}{5}$$

$$\boxed{\mathrm{P}(T = c) = P(S = r, T = c) + P(S = s, T = c)}$$

Conditional Probability

- Conditional probabilities are the ratio of two probabilities:

$$P(x_1|x_2) = \frac{P(x_1, x_2)}{P(x_2)}$$

$$P(x_1|x_2) = \frac{P(x_1, x_2)}{\sum_{x_1} P(x_1, x_2)}$$

Quiz: Conditional Probabilities

P(A, B)

A	B	P
+a	-b	0.5
+a	+b	0.1
-a	-b	0.2
-a	+b	0.2

$$P(a|b) = \frac{P(a,b)}{P(b)}$$

$$P(+a \mid +b) = ?$$

$$P(-a \mid +b) = ?$$

$$P(-b \mid +a) = ?$$

Quiz: Conditional Probabilities

P(A, B)

A	B	P
+a	-b	0.5
+a	+b	0.1
-a	-b	0.2
-a	+b	0.2

$$P(a|b) = \frac{P(a,b)}{P(b)}$$

$$P(+a \mid +b) = 1/3$$

$$P(-a \mid +b) = 2/3$$

$$P(-b \mid +a) = 5/6$$

Conditional Probability Representation

- P(T, S) = 2x2 table summing to 1
- P(warm | sun) = a single number
- P(T | S) = ?
  - Two 2-element vectors, each summing to 1
  - Conditional Distribution!

T	*S*	*P*
warm	sun	0.4
warm	rain	0.1
cold	sun	0.2
cold	rain	0.3

Conditional Distributions

- Conditional distributions are probability distributions over some variables *given* fixed values of others

Joint Distribution

$P(T, S)$

T	S	P
warm	sun	0.4
warm	rain	0.1
cold	sun	0.2
cold	rain	0.3

Conditional Distribution: P(T|S)

$P(T|S = r)$

T	P
warm	0.25
cold	0.75

$P(T|S = s)$

T	P
warm	0.67
cold	0.33

- How to derive conditional distributions from joint distributions?

Conditional Distributions: Normalization Trick

- A trick to get the conditional distribution based on two steps:
  1. Select the joint probabilities for each value of the conditioning variable
  2. Normalize the resulting vector

T	S	P
warm	sun	0.4
warm	rain	0.1
cold	sun	0.2
cold	rain	0.3

Select $\longrightarrow$

$P(T, r)$

T	P
warm	0.1
cold	0.3

Normalize $\longrightarrow$

$P(T|r)$

T	P
warm	0.25
cold	0.75

Conditional Distributions: Normalization Trick

T	S	P
warm	sun	0.4
warm	rain	0.1
cold	sun	0.2
cold	rain	0.3

Select $\longrightarrow$

$P(T, r)$

T	P
warm	0.1
cold	0.3

Normalize $\longrightarrow$

$P(T|r)$

T	P
warm	0.25
cold	0.75

- Why does this work?
  - The sum of the selection is: P(r), which is the probability of evidence!

$$P(x_1|x_2) = \frac{P(x_1, x_2)}{P(x_2)} = \frac{P(x_1, x_2)}{\sum_{x_1} P(x_1, x_2)}$$

To Normalize: Definition

- Definition: To make all the values in a table sum to one
  - Step 1: Compute the sum (Z) of all entries in the table
  - Step 2: Divide each entry by the sum

$P(T, S = s)$

T	P
warm	0.4
cold	0.2

$\xrightarrow[Z=0.6]{\text{Normalize}}$

T	P
warm	2/3
cold	1/3

Probabilistic Inference

- Compute a desired probability from other known probabilities
  - For example, conditional from joint

- We generally compute conditional probabilities that helps us summarize the beliefs given evidence!
  - P(A_{15} | no reported accidents) = 0.04
  - Probabilities change with new evidence:
    - P(A_{15} | no reported accidents, 5 a.m.) = 0.15
    - P(A_{15} | no reported accidents, 5 a.m., raining) = 0.08
      – i.e., observing evidence causes *beliefs to be updated*

Inference by Enumeration

- We can answer any query given a joint distribution!

- $\mathrm{P}(R = sun)$?

  Query

S	T	R	P
summer	warm	sun	0.30
summer	warm	rain	0.05
summer	cold	sun	0.10
summer	cold	rain	0.05
winter	warm	sun	0.10
winter	warm	rain	0.05
winter	cold	sun	0.15
winter	cold	rain	0.20

Inference by Enumeration

- We can answer any query given a joint distribution!

- $\mathrm{P}(R = sun)$ =0.3+0.1+0.1+0.15 = 0.65

  Query

S	T	R	P
summer	warm	sun	0.30
summer	warm	rain	0.05
summer	cold	sun	0.10
summer	cold	rain	0.05
winter	warm	sun	0.10
winter	warm	rain	0.05
winter	cold	sun	0.15
winter	cold	rain	0.20

Inference by Enumeration

- We can answer any query given a joint distribution!

- $\mathrm{P}(R = sun) =0.3+0.1+0.1+0.15 = 0.65$

  Query

- $\mathrm{P}(R = rain) =?$

S	T	R	P
summer	warm	sun	0.30
summer	warm	rain	0.05
summer	cold	sun	0.10
summer	cold	rain	0.05
winter	warm	sun	0.10
winter	warm	rain	0.05
winter	cold	sun	0.15
winter	cold	rain	0.20

Inference by Enumeration

- We can answer any query given a joint distribution!

- $\mathrm{P}(R = sun)$ =0.3+0.1+0.1+0.15 = 0.65

  Query

- $\mathrm{P}(R = rain) = 0.05+0.05+0.05+0.20 = 0.35$

S	T	R	P
summer	warm	sun	0.30
summer	warm	rain	0.05
summer	cold	sun	0.10
summer	cold	rain	0.05
winter	warm	sun	0.10
winter	warm	rain	0.05
winter	cold	sun	0.15
winter	cold	rain	0.20

Practice: Inference by Enumeration

- We can answer any query given a joint distribution!

- $P(R \mid \text{winter, warm}) = ?$

Query Evidence

S	*T*	*R*	*P*
summer	warm	sun	0.30
summer	warm	rain	0.05
summer	cold	sun	0.10
summer	cold	rain	0.05
winter	warm	sun	0.10
winter	warm	rain	0.05
winter	cold	sun	0.15
winter	cold	rain	0.20

1. Unnormalized sum of the entries consistent with evidence:

$$P(R=r, S=\text{winter}, T=\text{warm}) = 0.05$$
$$P(R=s, S=\text{winter}, T=\text{warm}) = 0.10$$

2. Normalize:

$$P(R=r \mid \text{winter, warm}) = 1/3\ ,\ \ P(R=s \mid \text{winter, warm}) = 2/3$$

Practice: Inference by Enumeration

- P(R | winter) = ?

Query Evidence T: Hidden or unobserved

Unnormalized sum of the entries consistent with evidence:

P(R= r, S=winter) = 0.25

P(R= s, S=winter) = 0.25

S	*T*	*R*	*P*
summer	warm	sun	0.30
summer	warm	rain	0.05
summer	cold	sun	0.10
summer	cold	rain	0.05
winter	warm	sun	0.10
winter	warm	rain	0.05
winter	cold	sun	0.15
winter	cold	rain	0.20

In the process, we summed out (marginalization) hidden variable T

P(R= r, S=winter) = P(R= r, T=w, S=winter) + P(R=r, T=c, S= winter) = 0.25

P(R= s, S=winter) = P(R= s, T=w, S= winter) + P(R=s, T=c, S= winter) = 0.25

Practice: Inference by Enumeration

- P(R | winter) = ?

Query Evidence T: hidden or unobserved

1. Unnormalized sum of the entries consistent with evidence:

Summing out T: P(R= r, S=winter) = 0.25

Summing out T: P(R= s, S=winter) = 0.25

2. Normalize:

Normalized: P(R= r | winter) = 0.5

Normalized: P(R= s | winter) = 0.5

S	*T*	*R*	*P*
summer	warm	sun	0.30
summer	warm	rain	0.05
summer	cold	sun	0.10
summer	cold	rain	0.05
winter	warm	sun	0.10
winter	warm	rain	0.05
winter	cold	sun	0.15
winter	cold	rain	0.20

Inference by Enumeration

- General case: All variables: $X_1, X_2, \dots X_n$
  - Evidence variables: $(E_1 \dots E_k) = (e_1 \dots e_k)$
  - Query variables: $Y_1 \dots Y_m$
  - Hidden variables: $H_1 \dots H_r$

- We want: $P(Y_1 \dots Y_m | e_1 \dots e_k)$
  - Select the entries consistent with the evidence and sum out **H**:

$$P(Y_1 \dots Y_m, e_1 \dots e_k) = \sum_{h_1 \dots h_r} P(Y_1 \dots Y_m, h_1 \dots h_r, e_1 \dots e_k)$$

  - Normalize the final values to obtain the probability

Inference by Enumeration: Complexity

- Space Complexity: $O(d^n)$ to store the joint distribution
- Worst case time complexity: $O(d^n)$
- Probabilistic inference is of exponential complexity!
  - What is probabilistic inference any way?
    - Predictive or Causal analytics
- How to make this problem tractable?

S	T	R	P
summer	warm	sun	0.30
summer	warm	rain	0.05
summer	cold	sun	0.10
summer	cold	rain	0.05
winter	warm	sun	0.10
winter	warm	rain	0.05
winter	cold	sun	0.15
winter	cold	rain	0.20

Thank You