L11 Uncertainty

Random Variable

An **aspect of the problem domain** which we **may have uncertainty** about

Random Variable (RV)	Description	Domain (D)
D	Result from flipping a die	{1,2,3,4,5,6}
R	Raining?	{yes, no}
M	Winning a football match	{win, lose}

Explanation

Random Variable : Always capital letters

Domain : Small letters enclosed in `{}`; A list of possible values

Probability Distribution Table

Specifies the $P()$ of all the values (outcomes) for a given **RV**

Example

T	P(T)
hot	0.6
cold	0.4

Probability Rules

- $\forall_x$ within the **RV**, $0 \leq P(x) \leq 1$
- Sum of all $P(x)$ is 1

Joint Distribution

A table specifying the $P()$ for multiple **RV**s

Probalistic Model : joint distribution over a set of **RV**s

W = {*hazy, sunny, rainy*}
T = {*hot, cold*}

$P(T, W)$

	hazy	*sunny*	*rainy*
hot	0.25	0.3	0.05
cold	0.15	0.05	0.20

$P(T, W)$

hot, hazy	0.25
cold, hazy	0.15
hot, sunny	0.30
cold, sunny	0.05
hot, rainy	0.05
cold, rainy	0.20

Note : The probability rules apply here too

Events

It is a **set** of outcomes for the **set of RV** (inferred from Joint Distributions)

Calculating Probability of an Event

Add up all the entries which are consistent with the event

$P(E) = \sum_{x_1, x_2, \dots, x_n} P(x_1, x_2, \dots, x_n)$

Example : Find the probability of hazy

$P(T, W)$

	hazy	*sunny*	*rainy*
hot	0.25	0.3	0.05
cold	0.15	0.05	0.20

P(*hazy*) = P(hazy, hot) + P(hazy, cold)
= 0.25 + 0.15 = 0.40

`^` AND, `v` OR

Example Events

$P(T, W)$

	hazy	*sunny*	*rainy*
hot	0.25	0.3	0.05
cold	0.15	0.05	0.20

- P(hot)? 0.25 + 0.3 + 0.05 = 0.6
- $P(\neg hazy)$? 0.3 + 0.05 + 0.05 + 0.2 = 0.6
- $P(\neg hazy \lor cold)$? 0.3 + 0.05 + 0.05 + 0.2 + 0.15 = 0.75
- P(hot $\land$ $\neg$sunny)? 0.25 + 0.05 = 0.30
- $P(hazy \land rainy)$? Not possible

Marginal Distributions

Sub-tables which **eliminates** certain variables

The process of **summing out**/**marginalization** helps us extract the distribution over some subset of variables

Full joint distribution
P(T, W)

	hazy	*sunny*	*rainy*
hot	0.25	0.3	0.05
cold	0.15	0.05	0.20

sum out W

Marginalized
Distribution
P(T)

hot	0.6
cold	0.4

P(hot) = P(hot, hazy) + P(hot, sunny) + P(hot, rainy)
P(cold) = P(cold, hazy) + P(cold, sunny) + P(cold, rainy)

sum out T

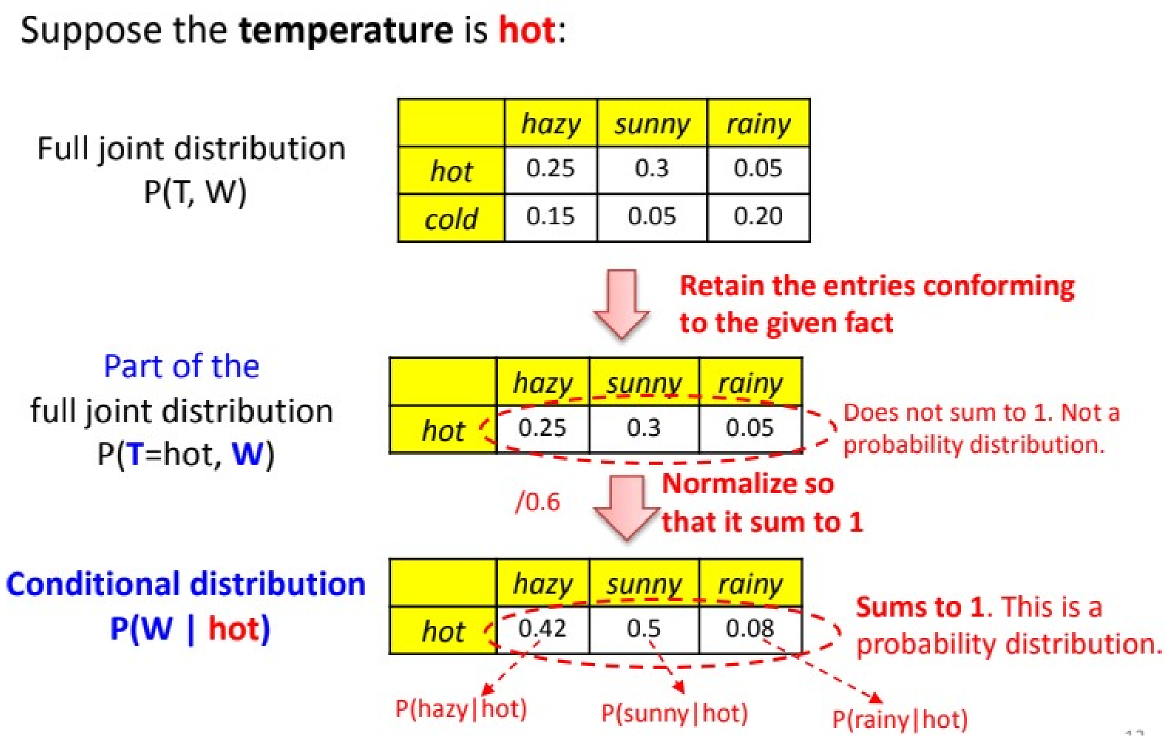
Marginalized
Distribution
P(W)

hazy	*sunny*	*rainy*
0.4	0.35	0.25

Conditional Distributions

$P(W|T)$ is the probability of the weather given it is hot/cold

Example : Converting a joint distribution $P(W, T)$ to conditional



Product Rule

$P(x,y) = P(x|y)P(y) \equiv P(x|y) = \frac{P(x,y)}{P(y)}$

Using that,

- Example:

	hazy	sunny	rainy
hot	0.25	0.3	0.05
cold	0.15	0.05	0.20

P(hazy) = P(hazy, hot) + P(hazy, cold)
= 0.25 + 0.15 = 0.4

P(hot | hazy) = P(hot, hazy) / P(hazy)
= 0.25 / 0.4
= 0.625

P(hazy, hot) = P(hot | hazy) P(hazy)
= 0.625×0.4
= 0.25

Refer to Slide 17,18 Example 2 and 3

Probabilistic Inference

Compute a desired probability from other know probas (e.g conditional from joint)

Inference by enumeration

Given the following:

- Query event: Q_E
- Query variables: $Q = Q_1 \dots Q_p$
- Evidence variables: $E = E_1 \dots E_k = e_1 \dots e_k$
- Hidden variables: $H = H_1 \dots H_r$

and we want to find:

$P(Q_E \mid e_1 \dots e_k)$

The inference process:

1. Select consistent entries
Select the entries consistent with the evidence to get: $P(Q, H, e_1 \dots e_k)$

2. Marginalize
Sum out H to get the joint distribution of query and evidence to get: $P(Q, e_1 \dots e_k)$

3. Normalize
Normalize to get the conditional probability: $P(Q \mid e_1 \dots e_k)$

4. Get Event
Sum the entries consistent with the event: $P(Q_E \mid e_1 \dots e_k)$

2

Causal and Diagnostic Probability

Take Cause = Dengue, Effect = Fever

Usually we have $P(\text{Effect}|\text{Cause})$, **causal probability**

We want this $P(\text{Cause}|\text{Effect})$, **diagnostic probability**

So we need **Bayes Rule** to help us get the reverse

$P(x|y) = \frac{P(y|x)P(x)}{P(y)}$ OR $P(\text{Cause}|\text{Effect}) = \frac{P(Effect|Cause)P(Cause)}{P(Effect)}$