# PROBABILITY AND INFERENCE

## Progress Report

- We've finished Part I: Problem Solving!
- Part II: Reasoning with uncertainty
- Part III: Machine Learning

## Today

- Random variables and probabilities
- Joint, marginal, conditional distributions
- Inference
- (Leading to Independence)

We're going to be using these concepts a lot so it's worth learning it well the first time!

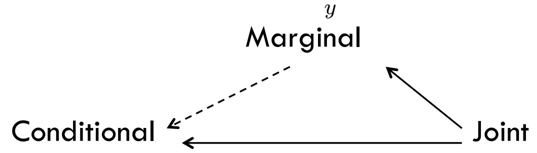
## Handling Uncertainty

- The world is an uncertain place
  - Partially observable, non-deterministic
  - On the way to the bank, you get in a car crash!
  - Medical diagnosis
  - Driving to Seattle
  - Sensors
- Probability theory gives us a language to reason about an uncertain world.
- Probability theory is beautiful!

## Random variables

## Discrete Probability distribution

# Distributions of Interest to Us




$$p(x) = \sum_y p(x, y)$$

$$p(x, y)$$

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

# Joint probability distribution

P(W = w, T = t)

| w | t | P |
|---|---|---|
| sunny | hot | 0.05 |
| rainy | hot | 0.01 |
| cloudy | hot | 0.01 |
| snowy | hot | 0.01 |
| sunny | cold | 0.20 |
| rainy | cold | 0.29 |
| cloudy | cold | 0.29 |
| snowy | cold | 0.14 |

# Events

P(W = w, T = t)

| w | t | P |
|---|---|---|
| sunny | hot | 0.07 |
| rainy | hot | 0.01 |
| cloudy | hot | 0.01 |
| snowy | hot | 0.01 |
| sunny | cold | 0.15 |
| rainy | cold | 0.3 |
| cloudy | cold | 0.3 |
| snowy | cold | 0.15 |

# Marginal Distributions

P(W = w, T = t)

| w | t | P |
|---|---|---|
| sunny | hot | 0.07 |
| rainy | hot | 0.01 |
| cloudy | hot | 0.01 |
| snowy | hot | 0.01 |
| sunny | cold | 0.15 |
| rainy | cold | 0.3 |
| cloudy | cold | 0.3 |
| snowy | cold | 0.15 |

# Conditional (posterior) distribution

- Often, we observe some information (evidence) and we want to know the probability of an event conditioned on this evidence

p(W | T = cold)

evidence

In all the worlds where T=cold, what is the probability that W = sunny? That W = rainy? That W = cloudy? That W = snowy?

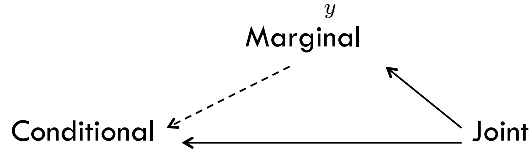
- This is called the conditional distribution, e.g. the distribution of W conditioned on the evidence T = cold

# Conditional (posterior) distribution

## Conditional and Joint are just a constant apart!

## Normalization Trick

# Distributions of Interest to Us




$$p(x) = \sum_y p(x, y)$$

$$p(x, y)$$

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

# Probabilistic Inference

- Probabilistic inference refers to the task of computing some desired probability given other known probabilities (evidence)

- Typically compute the conditional (posterior) probability of an event
  - p(on time to airport | no accidents ) = 0.80

- Probabilities change with new evidence
  - p(on time to airport | no accidents, 5 a.m.) = 0.95
  - p(on time to airport | no accidents, 5 a.m., raining) = 0.8

# Inference by Enumeration

# Inference by Enumeration

Step One: select the entries in the table consistent with the evidence (this becomes our world)

Step Two: sum over the H variables to get the joint distribution of the query and evidence variables

Step Three: Normalize

p(W | S = winter)?

| S | W | T | P |
|---|---|---|---|
| summer | sunny | hot | 0.30 |
| summer | rain | hot | 0.05 |
| summer | sunny | cold | 0.10 |
| summer | rain | cold | 0.05 |
| winter | sunny | hot | 0.10 |
| winter | rain | hot | 0.05 |
| winter | sunny | cold | 0.15 |
| winter | rain | cold | 0.20 |

Step One

| S | W | T | P |
|---|---|---|---|
| ~~summer~~ | ~~sunny~~ | ~~hot~~ | ~~0.30~~ |
| ~~summer~~ | ~~rain~~ | ~~hot~~ | ~~0.05~~ |
| ~~summer~~ | ~~sunny~~ | ~~cold~~ | ~~0.10~~ |
| ~~summer~~ | ~~rain~~ | ~~cold~~ | ~~0.05~~ |
| winter | sunny | hot | 0.10 |
| winter | rain | hot | 0.05 |
| winter | sunny | cold | 0.15 |
| winter | rain | cold | 0.20 |

# Inference by Enumeration

Step One: select the entries in the table consistent with the evidence (this becomes our world)

Step Two: sum over the H variables to get the joint distribution of the query and evidence variables

Step Three: Normalize

p(W | S = winter)?

| S | W | T | P |
|---|---|---|---|
| winter | sunny | hot | 0.10 |
| winter | rain | hot | 0.05 |
| winter | sunny | cold | 0.15 |
| winter | rain | cold | 0.20 |

Step Two

$$p(Q, e_1, \ldots, e_k) = \sum_{(h_1, \ldots, h_r)} p(Q, e_1, \ldots, e_k, h_1, \ldots, h_r)$$

# Inference by Enumeration

Step One: select the entries in the table consistent with the evidence (this becomes our world)

Step Two: sum over the H variables to get the joint distribution of the query and evidence variables

Step Three: Normalize

p(W | S = winter)?

| S | W | T | P |
|---|---|---|---|
| winter | sunny | hot | 0.10 |
| winter | rain | hot | 0.05 |
| winter | sunny | cold | 0.15 |
| winter | rain | cold | 0.20 |

Step Three

$$Z = \sum_q p(Q = q, e_1, \ldots, e_k)$$

$$p(Q | e_1, \ldots, e_k) = \frac{1}{Z} \cdot p(Q, e_1, \ldots, e_k)$$

# Inference by Enumeration

Step One: select the entries in the table consistent with the evidence (this becomes our world)

Step Two: sum over the H variables to get the joint distribution of the query and evidence variables

Step Three: Normalize

| S | W | T | P |
|---|---|---|---|
| summer | sunny | hot | 0.30 |
| summer | rain | hot | 0.05 |
| summer | sunny | cold | 0.10 |
| summer | rain | cold | 0.05 |
| winter | sunny | hot | 0.10 |
| winter | rain | hot | 0.05 |
| winter | sunny | cold | 0.15 |
| winter | rain | cold | 0.20 |

Queries:

P(S | T = hot)?

p(W | S=winter, T=hot)?

p(S, W)?

p(S, W | T=hot)?

# Inference by Enumeration

- n random variables
- d domain size
- Worst-case time is $O(d^n)$
- Space is $O(d^n)$ to save entire table in memory

- Is there something better?