

Random Variables

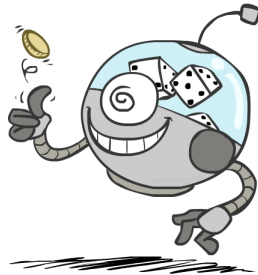
- A random variable is some aspect of the world about which we (may) have uncertainty

- R = Is it raining?
- T = Is it hot or cold?
- D = How long will it take to drive to work?
- L = Where is the ghost?

- We denote random variables with capital letters

- Like variables in a CSP, random variables have domains

- R in {true, false} (often write as {+r, -r})
- T in {hot, cold}
- D in [0, ∞)
- L in possible locations, maybe {(0,0), (0,1), ...}



Probability Distributions

- Associate a probability with each value

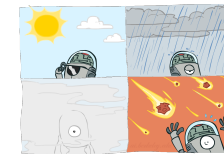
- Temperature:



$$P(T)$$

T	P
hot	0.5
cold	0.5

- Weather:



$$P(W)$$

W	P
sun	0.6
rain	0.1
fog	0.3
meteor	0.0

Probability Distributions

- Unobserved random variables have distributions

$$P(T)$$

T	P
hot	0.5
cold	0.5

$$P(W)$$

W	P
sun	0.6
rain	0.1
fog	0.3
meteor	0.0

Shorthand notation:

$$P(\text{hot}) = P(T = \text{hot}),$$

$$P(\text{cold}) = P(T = \text{cold}),$$

$$P(\text{rain}) = P(W = \text{rain}),$$

$$\dots$$

OK if all domain entries are unique

- A distribution is a TABLE of probabilities of values

- A probability (lower case value) is a single number

$$P(W = \text{rain}) = 0.1$$

- Must have: $\forall x \ P(X = x) \geq 0$ and $\sum_x P(X = x) = 1$

Joint Distributions

- A joint distribution over a set of random variables: X_1, X_2, \dots, X_n specifies a real number for each assignment (or outcome):

$$P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$$

$$P(x_1, x_2, \dots, x_n)$$

- Must obey: $P(x_1, x_2, \dots, x_n) \geq 0$

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

$P(T, W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

- Size of distribution if n variables with domain sizes d?

- For all but the smallest distributions, impractical to write out!

Probabilistic Models

- A probabilistic model is a joint distribution over a set of random variables
- Probabilistic models:
 - (Random) variables with domains
 - Assignments are called *outcomes*
 - Joint distributions: say whether assignments (outcomes) are likely
 - Normalized: sum to 1.0
 - Ideally: only certain variables directly interact
- Constraint satisfaction problems:
 - Variables with domains
 - Constraints: state whether assignments are possible
 - Ideally: only certain variables directly interact

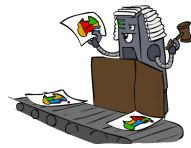
Distribution over T,W

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3



Constraint over T,W

T	W	P
hot	sun	T
hot	rain	F
cold	sun	F
cold	rain	T



Events

- An *event* is a set E of outcomes

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

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

- Probability that it's hot AND sunny?
- Probability that it's hot?
- Probability that it's hot OR sunny?

- Typically, the events we care about are *partial assignments*, like $P(T=\text{hot})$

$P(T, W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

Quiz: Events

- $P(+x, +y)$?
- $P(+x)$?
- $P(-y \text{ OR } +x)$?

$P(X, Y)$

X	Y	P
+x	+y	0.2
+x	-y	0.3
-x	+y	0.4
-x	-y	0.1

Marginal Distributions

- Marginal distributions are sub-tables which eliminate variables
- Marginalization (summing out): Combine collapsed rows by adding

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

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

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

T	P
hot	0.5
cold	0.5

W	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: Marginal Distributions

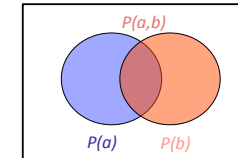
$P(X, Y)$				
X	Y	P		
+x	+y	0.2	$P(x) = \sum_y P(x, y)$	$P(X)$
+x	-y	0.3		
-x	+y	0.4	$P(y) = \sum_x P(x, y)$	$P(Y)$
-x	-y	0.1		



Conditional Probabilities

- A simple relation between joint and conditional probabilities
 - In fact, this is taken as the *definition* of a conditional probability

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



$P(T, W)$		
T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

$$P(W = s | T = c) = \frac{P(W = s, T = c)}{P(T = c)} = \frac{0.2}{0.5} = 0.4$$

$$= P(W = s, T = c) + P(W = r, T = c)$$

$$= 0.2 + 0.3 = 0.5$$

Quiz: Conditional Probabilities

$P(X, Y)$		
X	Y	P
+x	+y	0.2
+x	-y	0.3
-x	+y	0.4
-x	-y	0.1

- $P(+x | +y)$?
- $P(-x | +y)$?
- $P(-y | +x)$?

Conditional Distributions

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

Conditional Distributions

$P(W|T)$

$P(W|T = \text{hot})$

W	P
sun	0.8
rain	0.2

$P(W|T = \text{cold})$

W	P
sun	0.4
rain	0.6

Joint Distribution

$P(T, W)$		
T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

Normalization Trick

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

$$P(W = s|T = c) = \frac{P(W = s, T = c)}{P(T = c)}$$

$$= \frac{P(W = s, T = c)}{P(W = s, T = c) + P(W = r, T = c)}$$

$$= \frac{0.2}{0.2 + 0.3} = 0.4$$

W	P
sun	0.4
rain	0.6

$$P(W = r|T = c) = \frac{P(W = r, T = c)}{P(T = c)}$$

$$= \frac{P(W = r, T = c)}{P(W = s, T = c) + P(W = r, T = c)}$$

$$= \frac{0.3}{0.2 + 0.3} = 0.6$$

Normalization Trick

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

$$P(W = s|T = c) = \frac{P(W = s, T = c)}{P(T = c)}$$

$$= \frac{P(W = s, T = c)}{P(W = s, T = c) + P(W = r, T = c)}$$

$$= \frac{0.2}{0.2 + 0.3} = 0.4$$

W	P
sun	0.4
rain	0.6

$$P(W = r|T = c) = \frac{P(W = r, T = c)}{P(T = c)}$$

$$= \frac{P(W = r, T = c)}{P(W = s, T = c) + P(W = r, T = c)}$$

$$= \frac{0.3}{0.2 + 0.3} = 0.6$$

Normalization Trick

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

SELECT the joint probabilities matching the evidence
 \rightarrow

T	W	P
cold	sun	0.2
cold	rain	0.3

NORMALIZE the selection (make it sum to one)
 \rightarrow

W	P
sun	0.4
rain	0.6

- Why does this work? Sum of selection is $P(\text{evidence})!$ ($P(T=c)$, here)

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

Quiz: Normalization Trick

- $P(X | Y=-y)$?

X	Y	P
+x	+y	0.2
+x	-y	0.3
-x	+y	0.4
-x	-y	0.1

SELECT the joint probabilities matching the evidence
 \rightarrow

NORMALIZE the selection (make it sum to one)
 \rightarrow

To Normalize

- (Dictionary) To bring or restore to a **normal condition**

All entries sum to ONE

- Procedure:

- Step 1: Compute Z = sum over all entries
- Step 2: Divide every entry by Z

- Example 1

W	P
sun	0.2
rain	0.3

Normalize
 $Z = 0.5$

W	P
sun	0.4
rain	0.6

- Example 2

T	W	P
hot	sun	20
hot	rain	5
cold	sun	10
cold	rain	15

Normalize
 $Z = 50$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

The Product Rule

- Sometimes have conditional distributions but want the joint

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



The Product Rule

$$P(y)P(x|y) = P(x, y)$$

- Example:

$P(W)$		$P(D W)$			\longleftrightarrow	$P(D, W)$		
R	P	D	W	P		D	W	P
sun	0.8	wet	sun	0.1		wet	sun	
rain	0.2	dry	sun	0.9		dry	sun	
		wet	rain	0.7		wet	rain	
		dry	rain	0.3		dry	rain	

The Chain Rule

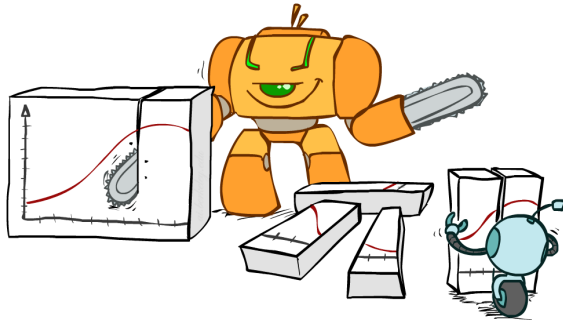
- More generally, can always write any joint distribution as an incremental product of conditional distributions

$$P(x_1, x_2, x_3) = P(x_1)P(x_2|x_1)P(x_3|x_1, x_2)$$

$$P(x_1, x_2, \dots, x_n) = \prod_i P(x_i|x_1 \dots x_{i-1})$$

- Why is this always true?

Bayes Rule



Bayes' Rule

- Two ways to factor a joint distribution over two variables:

$$P(x, y) = P(x|y)P(y) = P(y|x)P(x)$$

- Dividing, we get:

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

- Why is this at all helpful?

- Lets us build one conditional from its reverse
- Often one conditional is tricky but the other one is simple

- In the running for most important AI equation!

That's my rule!



Inference with Bayes' Rule

- Example: Diagnostic probability from causal probability:

$$P(\text{cause}|\text{effect}) = \frac{P(\text{effect}|\text{cause})P(\text{cause})}{P(\text{effect})}$$

- Example:

- M: meningitis, S: stiff neck

$$\left. \begin{array}{l} P(+m) = 0.0001 \\ P(+s | +m) = 0.8 \\ P(+s | -m) = 0.01 \end{array} \right\} \text{Example givens}$$

$$P(+m | +s) = \frac{P(+s | +m)P(+m)}{P(+s)} = \frac{P(+s | +m)P(+m)}{P(+s | +m)P(+m) + P(+s | -m)P(-m)} = \frac{0.8 \times 0.0001}{0.8 \times 0.0001 + 0.01 \times 0.999}$$

- Note: posterior probability of meningitis still very small
- Note: you should still get stiff necks checked out! Why?

Quiz: Bayes' Rule

- Given:

$$P(W)$$

R	P
sun	0.8
rain	0.2

$$P(D|W)$$

D	W	P
wet	sun	0.1
dry	sun	0.9
wet	rain	0.7
dry	rain	0.3

- What is $P(W | \text{dry})$?