CS 343: Artificial Intelligence

Probability

Prof. Scott Niekum — The University of Texas at Austin

[These slides based on those of Dan Klein and Pieter Abbeel for CS188 Intro to AI at UC Berkeley. All CS188 materials are available at http://ai.berkeley.edu.]
Today

- Probability
  - Random Variables
  - Joint and Marginal Distributions
  - Conditional Distributions
  - Product Rule, Chain Rule, Bayes’ Rule
  - Inference
  - Independence

- You’ll need all this stuff A LOT for the next few weeks, so make sure you go over it now!
Inference in Ghostbusters

- A ghost is in the grid somewhere
- Noisy sensor readings tell how close a square is to the ghost.

Most likely observations:
- On the ghost: red
- 1 or 2 away: orange
- 3 or 4 away: yellow
- 5+ away: green

- Sensors are noisy, but we know \( P(\text{Color} \mid \text{Distance}) \)

<table>
<thead>
<tr>
<th></th>
<th>( P(\text{red} \mid 3) )</th>
<th>( P(\text{orange} \mid 3) )</th>
<th>( P(\text{yellow} \mid 3) )</th>
<th>( P(\text{green} \mid 3) )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.05</td>
<td>0.15</td>
<td>0.5</td>
<td>0.3</td>
</tr>
</tbody>
</table>
Ghostbusters, no probabilities
Uncertainty

- General situation:
  - **Observed variables (evidence):** Agent knows certain things about the state of the world (e.g., sensor readings or symptoms)
  - **Unobserved variables:** Agent needs to reason about other aspects (e.g., where an object is or what disease is present)
  - **Model:** Agent knows something about how the known variables relate to the unknown variables

- Probabilistic reasoning gives us a framework for using beliefs and knowledge to perform inference
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 \([0, \infty)\)
  - \( L \) 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

<table>
<thead>
<tr>
<th></th>
<th>$P(T)$</th>
<th>$P(W)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>P</td>
<td>W</td>
</tr>
<tr>
<td>hot</td>
<td>0.5</td>
<td>sun</td>
</tr>
<tr>
<td>cold</td>
<td>0.5</td>
<td>rain</td>
</tr>
<tr>
<td>fog</td>
<td>0.3</td>
<td>meteor</td>
</tr>
</tbody>
</table>

- A discrete distribution is a table of probabilities of values
- A probability (lower case value) is a single number

$P(W = rain) = 0.1$

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

Shorthand notation:

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

OK if all domain entries are unique
Joint Distributions

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

\[
P(X_1 = x_1, X_2 = x_2, \ldots X_n = x_n) \\
P(x_1, x_2, \ldots x_n)
\]

- Must obey:

\[
P(x_1, x_2, \ldots x_n) \geq 0 \\
\sum_{(x_1,x_2,\ldots x_n)} P(x_1, x_2, \ldots x_n) = 1
\]

- Size of distribution if $n$ variables with domain sizes $d$?
  - For all but the smallest distributions, impractical to write out!
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:

<table>
<thead>
<tr>
<th>T</th>
<th>W</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>hot</td>
<td>sun</td>
<td>0.4</td>
</tr>
<tr>
<td>hot</td>
<td>rain</td>
<td>0.1</td>
</tr>
<tr>
<td>cold</td>
<td>sun</td>
<td>0.2</td>
</tr>
<tr>
<td>cold</td>
<td>rain</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Constraint over T,W:

<table>
<thead>
<tr>
<th>T</th>
<th>W</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>hot</td>
<td>sun</td>
<td>T</td>
</tr>
<tr>
<td>hot</td>
<td>rain</td>
<td>F</td>
</tr>
<tr>
<td>cold</td>
<td>sun</td>
<td>F</td>
</tr>
<tr>
<td>cold</td>
<td>rain</td>
<td>T</td>
</tr>
</tbody>
</table>
An event is a set $E$ of outcomes

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

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 $P(T=\text{hot})$
Quiz: Events

- P(+x, +y) ?
- P(+x) ?
- P(-y OR +x) ?

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>+x</td>
<td>+y</td>
<td>0.2</td>
</tr>
<tr>
<td>+x</td>
<td>-y</td>
<td>0.3</td>
</tr>
<tr>
<td>-x</td>
<td>+y</td>
<td>0.4</td>
</tr>
<tr>
<td>-x</td>
<td>-y</td>
<td>0.1</td>
</tr>
</tbody>
</table>
Marginal Distributions

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

\[
P(T, W)
\]

<table>
<thead>
<tr>
<th>T</th>
<th>W</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>hot</td>
<td>sun</td>
<td>0.4</td>
</tr>
<tr>
<td>hot</td>
<td>rain</td>
<td>0.1</td>
</tr>
<tr>
<td>cold</td>
<td>sun</td>
<td>0.2</td>
</tr>
<tr>
<td>cold</td>
<td>rain</td>
<td>0.3</td>
</tr>
</tbody>
</table>

\[
P(T) = \sum_s P(t, s)
\]

<table>
<thead>
<tr>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>hot</td>
<td>0.5</td>
</tr>
<tr>
<td>cold</td>
<td>0.5</td>
</tr>
</tbody>
</table>

\[
P(W) = \sum_t P(t, s)
\]

<table>
<thead>
<tr>
<th>W</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>sun</td>
<td>0.6</td>
</tr>
<tr>
<td>rain</td>
<td>0.4</td>
</tr>
</tbody>
</table>

\[
P(X_1 = x_1) = \sum_{x_2} P(X_1 = x_1, X_2 = x_2)
\]
Quiz: Marginal Distributions

\[
P(X, Y) = \begin{array}{ccc} 
  X & Y & P \\
  +x & +y & 0.1 \\
  +x & -y & 0.5 \\
  -x & +y & 0.2 \\
  -x & -y & 0.2 \\
\end{array}
\]

\[
P(x) = \sum_y P(x, y)
\]

\[
P(y) = \sum_x P(x, y)
\]
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>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>W</td>
<td>P</td>
</tr>
<tr>
<td>hot</td>
<td>sun</td>
<td>0.4</td>
</tr>
<tr>
<td>hot</td>
<td>rain</td>
<td>0.1</td>
</tr>
<tr>
<td>cold</td>
<td>sun</td>
<td>0.2</td>
</tr>
<tr>
<td>cold</td>
<td>rain</td>
<td>0.3</td>
</tr>
</tbody>
</table>

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

\[
= P(W = s, T = c) + P(W = r, T = c) = 0.2 + 0.3 = 0.5
\]
Quiz: Conditional Probabilities

\[ P(X, Y) \]

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>+x</td>
<td>+y</td>
<td>0.2</td>
</tr>
<tr>
<td>+x</td>
<td>-y</td>
<td>0.3</td>
</tr>
<tr>
<td>-x</td>
<td>+y</td>
<td>0.4</td>
</tr>
<tr>
<td>-x</td>
<td>-y</td>
<td>0.1</td>
</tr>
</tbody>
</table>

- \( P(+x \mid +y) \) ?
- \( P(-x \mid +y) \) ?
- \( P(-y \mid +x) \) ?
Conditional Distributions

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

### Conditional Distributions

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

<table>
<thead>
<tr>
<th>W</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>sun</td>
<td>0.8</td>
</tr>
<tr>
<td>rain</td>
<td>0.2</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>W</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>sun</td>
<td>0.4</td>
</tr>
<tr>
<td>rain</td>
<td>0.6</td>
</tr>
</tbody>
</table>

### Joint Distribution

\[
P(T, W)
\]

<table>
<thead>
<tr>
<th>T</th>
<th>W</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>hot</td>
<td>sun</td>
<td>0.4</td>
</tr>
<tr>
<td>hot</td>
<td>rain</td>
<td>0.1</td>
</tr>
<tr>
<td>cold</td>
<td>sun</td>
<td>0.2</td>
</tr>
<tr>
<td>cold</td>
<td>rain</td>
<td>0.3</td>
</tr>
</tbody>
</table>
Normalization Trick

\[ P(W = s | T = c) = \frac{P(W = s, T = c)}{P(T = c)} \]
\[ = \frac{P(W = s, T = c)}{P(W = s, T = c) + P(W = r, T = c)} \]
\[ = \frac{0.2}{0.2 + 0.3} = 0.4 \]

\[ P(W = r | T = c) = \frac{P(W = r, T = c)}{P(T = c)} \]
\[ = \frac{P(W = r, T = c)}{P(W = s, T = c) + P(W = r, T = c)} \]
\[ = \frac{0.3}{0.2 + 0.3} = 0.6 \]
Normalization Trick

\[ P(W = s | T = c) = \frac{P(W = s, T = c)}{P(T = c)} \]
\[ = \frac{P(W = s, T = c)}{P(W = s, T = c) + P(W = r, T = c)} \]
\[ = \frac{0.2}{0.2 + 0.3} = 0.4 \]

**P(T, W)**

<table>
<thead>
<tr>
<th>T</th>
<th>W</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>hot</td>
<td>sun</td>
<td>0.4</td>
</tr>
<tr>
<td>hot</td>
<td>rain</td>
<td>0.1</td>
</tr>
<tr>
<td>cold</td>
<td>sun</td>
<td>0.2</td>
</tr>
<tr>
<td>cold</td>
<td>rain</td>
<td>0.3</td>
</tr>
</tbody>
</table>

**SELECT** the joint probabilities matching the evidence

**P(c, W)**

<table>
<thead>
<tr>
<th>T</th>
<th>W</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>cold</td>
<td>sun</td>
<td>0.2</td>
</tr>
<tr>
<td>cold</td>
<td>rain</td>
<td>0.3</td>
</tr>
</tbody>
</table>

**NORMALIZE** the selection (make it sum to one)

\[ P(W | T = c) \]

<table>
<thead>
<tr>
<th>W</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>sun</td>
<td>0.4</td>
</tr>
<tr>
<td>rain</td>
<td>0.6</td>
</tr>
</tbody>
</table>

\[ P(W = r | T = c) = \frac{P(W = r, T = c)}{P(T = c)} \]
\[ = \frac{P(W = r, T = c)}{P(W = s, T = c) + P(W = r, T = c)} \]
\[ = \frac{0.3}{0.2 + 0.3} = 0.6 \]
### Normalization Trick

| $P(T, W)$ |  |
|---|---|---|
| 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

| $P(c, W)$ |  |
|---|---|---|
| T | W | P |
| cold | sun | 0.2 |
| cold | rain | 0.3 |

**NORMALIZE** the selection (make it sum to one)

<table>
<thead>
<tr>
<th>$P(W</th>
<th>T = c)$</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>sun</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>rain</td>
<td>0.6</td>
<td></td>
</tr>
</tbody>
</table>

Why does this work?

$$P(x_1|x_2) = \frac{P(x_1, x_2)}{P(x_2)} = \frac{P(x_1, x_2)}{\sum_{x_1} P(x_1, x_2)}$$
Quiz: Normalization Trick

- $P(X \mid Y=-y)$?

<table>
<thead>
<tr>
<th>$P(X, Y)$</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$X$</td>
<td>$Y$</td>
<td>$P$</td>
</tr>
<tr>
<td>$+x$</td>
<td>$+y$</td>
<td>0.3</td>
</tr>
<tr>
<td>$+x$</td>
<td>$-y$</td>
<td>0.1</td>
</tr>
<tr>
<td>$-x$</td>
<td>$+y$</td>
<td>0.5</td>
</tr>
<tr>
<td>$-x$</td>
<td>$-y$</td>
<td>0.1</td>
</tr>
</tbody>
</table>

**SELECT** the joint probabilities matching the evidence

**NORMALIZE** the selection (make it sum to one)
(Dictionary) To bring or restore to a normal condition

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

Example 1

<table>
<thead>
<tr>
<th>W</th>
<th>P</th>
<th>Normalize</th>
<th>W</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>sun</td>
<td>0.2</td>
<td>Z = 0.5</td>
<td>sun</td>
<td>0.4</td>
</tr>
<tr>
<td>rain</td>
<td>0.3</td>
<td></td>
<td>rain</td>
<td>0.6</td>
</tr>
</tbody>
</table>

Example 2

<table>
<thead>
<tr>
<th>T</th>
<th>W</th>
<th>P</th>
<th>Normalize</th>
<th>T</th>
<th>W</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>hot</td>
<td>sun</td>
<td>20</td>
<td></td>
<td>hot</td>
<td>sun</td>
<td>0.4</td>
</tr>
<tr>
<td>hot</td>
<td>rain</td>
<td>5</td>
<td></td>
<td>hot</td>
<td>rain</td>
<td>0.1</td>
</tr>
<tr>
<td>cold</td>
<td>sun</td>
<td>10</td>
<td></td>
<td>cold</td>
<td>sun</td>
<td>0.2</td>
</tr>
<tr>
<td>cold</td>
<td>rain</td>
<td>15</td>
<td></td>
<td>cold</td>
<td>rain</td>
<td>0.3</td>
</tr>
</tbody>
</table>
Probabilistic Inference

- Probabilistic inference: compute a desired probability from other known probabilities (e.g. conditional from joint)

- We generally compute conditional probabilities
  - P(on time | no reported accidents) = 0.90
  - These represent the agent’s beliefs given the evidence

- Probabilities change with new evidence:
  - P(on time | no accidents, 5 a.m.) = 0.95
  - P(on time | no accidents, 5 a.m., raining) = 0.80
  - Observing new evidence causes beliefs to be updated
Inference by Enumeration

- **General case:**
  - Evidence variables: $E_1 \ldots E_k = e_1 \ldots e_k$
  - Query* variable: $Q$
  - Hidden variables: $H_1 \ldots H_r$

\[
\begin{align*}
\{ & X_1, X_2, \ldots X_n \\
& \text{All variables} \}
\end{align*}
\]

- **Step 1:** Select the entries consistent with the evidence

- **Step 2:** Sum out $H$ to get joint of Query and evidence

- **Step 3:** Normalize

\[
P(Q|e_1 \ldots e_k) = \frac{1}{Z} \sum_{h_1 \ldots h_r} P(Q, h_1 \ldots h_r, e_1 \ldots e_k)
\]

\[
Z = \sum_{q} P(Q, e_1 \ldots e_k)
\]

* Works fine with multiple query variables, too

---

* Step 1: Select the entries consistent with the evidence

* Step 2: Sum out $H$ to get joint of Query and evidence

* Step 3: Normalize
Inference by Enumeration

- **P(W)?**
  - \( p(W=\text{sun}) = 0.3 + 0.1 + 0.1 + 0.15 = 0.65 \)
  - \( p(W=\text{rain}) = 0.05 + 0.05 + 0.05 + 0.2 = 0.35 \)

- **P(W | winter)?**
  - \( p(W=\text{sun}, \text{winter}) = 0.1 + 0.15 = 0.25 \)
  - \( p(W=\text{rain}, \text{winter}) = 0.05 + 0.2 = 0.25 \)
  - \( p(W=\text{sun} | \text{winter}) = 0.25 / 0.25 + 0.25 = 0.5 \)
  - \( p(W=\text{rain} | \text{winter}) = 0.25 / 0.25 + 0.25 = 0.5 \)

- **P(W | winter, hot)?**
  - \( p(W=\text{sun}, \text{winter, hot}) = 0.1 \)
  - \( p(W=\text{rain}, \text{winter, hot}) = 0.05 \)
  - \( p(W=\text{sun} | \text{winter, hot}) = 0.1 / 0.1 + 0.05 = 2/3 \)
  - \( p(W=\text{rain} | \text{winter, hot}) = 0.05 / 0.1 + 0.05 = 1/3 \)

<table>
<thead>
<tr>
<th>S</th>
<th>T</th>
<th>W</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>summer</td>
<td>hot</td>
<td>sun</td>
<td>0.30</td>
</tr>
<tr>
<td>summer</td>
<td>hot</td>
<td>rain</td>
<td>0.05</td>
</tr>
<tr>
<td>summer</td>
<td>cold</td>
<td>sun</td>
<td>0.10</td>
</tr>
<tr>
<td>summer</td>
<td>cold</td>
<td>rain</td>
<td>0.05</td>
</tr>
<tr>
<td>winter</td>
<td>hot</td>
<td>sun</td>
<td>0.10</td>
</tr>
<tr>
<td>winter</td>
<td>hot</td>
<td>rain</td>
<td>0.05</td>
</tr>
<tr>
<td>winter</td>
<td>cold</td>
<td>sun</td>
<td>0.15</td>
</tr>
<tr>
<td>winter</td>
<td>cold</td>
<td>rain</td>
<td>0.20</td>
</tr>
</tbody>
</table>
Inference by Enumeration

- **Obvious problems:**
  - Worst-case time complexity $O(d^n)$
  - Space complexity $O(d^n)$ to store the joint distribution
  - What about continuous distributions?
The Product Rule

- Sometimes have conditional distributions but want the joint

\[ P(y)P(x|y) = P(x, y) \quad \iff \quad P(x|y) = \frac{P(x, y)}{P(y)} \]
The Product Rule

\[ P(y) P(x|y) = P(x, y) \]

Example:

\[
\begin{array}{ccc}
\text{P}(\text{W}) & \text{D} & \text{W} & \text{P} \\
\text{R} & \text{P} & \text{wet} & \text{sun} & 0.1 \\
\text{sun} & 0.8 & \text{dry} & \text{sun} & 0.9 \\
\text{rain} & 0.2 & \text{wet} & \text{rain} & 0.7 \\
\text{dry} & \text{rain} & 0.3 & \\
\end{array}
\]

\[
\begin{array}{ccc}
\text{P}(\text{D}, \text{W}) & \text{D} & \text{W} & \text{P} \\
\text{wet} & \text{sun} & \\
\text{dry} & \text{sun} & \\
\text{wet} & \text{rain} & \\
\text{dry} & \text{rain} & \\
\end{array}
\]
The Chain Rule

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

\[ P(x_1, x_2, x_3) = P(x_1)P(x_2|x_1)P(x_3|x_1, x_2) \]

\[ P(x_1, x_2, \ldots x_n) = \prod_{i} P(x_i|x_1 \ldots x_{i-1}) \]

- Why is this always true?
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
  - Foundation of many systems we’ll see later

- In the running for most important AI equation!
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

\[
\begin{align*}
P(+m) &= 0.0001 \\
P(+s|m) &= 0.8 \\
P(+s|m) &= 0.01
\end{align*}
\]

\[
P(+m|+s) = \frac{P(+s|m)P(+m)}{P(+s)} = \frac{P(+s|m)P(+m)}{P(+s|m)P(+m) + P(+s|m)P(-m)P(-m)} = \frac{0.8 \times 0.0001}{0.8 \times 0.0001 + 0.01 \times 0.999} = 0.0008
\]

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

- **Given:**

  \[ P(W) \]
  \[
  \begin{array}{c|c}
  \hline
  \text{R} & \text{P} \\
  \hline
  \text{sun} & 0.8 \\
  \text{rain} & 0.2 \\
  \hline
  \end{array}
  \]

  \[ P(D|W) \]
  \[
  \begin{array}{c|c|c}
  \hline
  \text{D} & \text{W} & \text{P} \\
  \hline
  \text{wet} & \text{sun} & 0.1 \\
  \text{dry} & \text{sun} & 0.9 \\
  \text{wet} & \text{rain} & 0.7 \\
  \text{dry} & \text{rain} & 0.3 \\
  \hline
  \end{array}
  \]

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

  \[ p(\text{sun} | \text{dry}) = p(\text{dry} | \text{sun}) \frac{p(\text{sun})}{p(\text{dry})} = 0.9 \times 0.8 / Z = .72 / Z \]

  \[ p(\text{rain} | \text{dry}) = p(\text{dry} | \text{rain}) \frac{p(\text{rain})}{p(\text{dry})} = 0.3 \times 0.2 / Z = 0.06 / Z \]

  \[ Z = .72 + .06 = .78 \]
Let’s say we have two distributions:
- Prior distribution over ghost location: $P(G)$
  - Let’s say this is uniform
- Sensor reading model: $P(R \mid G)$
  - Given: we know what our sensors do
  - $R$ = reading color measured at (1,1)
  - E.g. $P(R = \text{yellow} \mid G=(1,1)) = 0.1$

We can calculate the posterior distribution $P(G \mid r)$ over ghost locations given a reading using Bayes’ rule:

$$P(g \mid r) \propto P(r \mid g)P(g)$$
Ghostbusters with Probability