Quantifying Uncertainty

Dr. Steven Bethard

Computer and Information Sciences
University of Alabama at Birmingham

8 Mar 2016

Outline

Outline

Motivation

Goal

Look out window - is it warm?

Rules

$\forall d\ Sunny(d) \Rightarrow Warm(d)$

Wrong; it can be sunny but cold

$\forall d\ Warm(d) \Rightarrow Sunny(d)$

Wrong; it can be warm and cloudy

Problem

Can't capture that sunny days are usually warm

Motivation

Goal

Look out window - is it warm?

Rules

$\forall d\ Sunny(d) \Rightarrow Warm(d)$

Wrong; it can be sunny but cold

$\forall d\ Warm(d) \Rightarrow Sunny(d)$

Wrong; it can be warm and cloudy

Problem

Can't capture that sunny days are usually warm

Motivation

Goal

Look out window - is it warm?

Rules

$\forall d \; Sunny(d) \Rightarrow Warm(d)$

Wrong; it can be sunny but cold

$\forall d \; Warm(d) \Rightarrow Sunny(d)$

Wrong; it can be warm and cloudy

Problem

Can't capture that sunny days are usually warm

Motivation

Goal

Look out window - is it warm?

Rules

$\forall d\ Sunny(d) \Rightarrow Warm(d)$

Wrong; it can be sunny but cold

$\forall d\ Warm(d) \Rightarrow Sunny(d)$

Wrong; it can be warm and cloudy

Problem

Can't capture that sunny days are usually warm

Motivation

Goal

Look out window - is it warm?

Rules

$\forall d \; Sunny(d) \Rightarrow Warm(d)$

Wrong; it can be sunny but cold

$\forall d \; Warm(d) \Rightarrow Sunny(d)$

Wrong; it can be warm and cloudy

Problem

Can't capture that sunny days are usually warm

Motivation

Goal

Look out window - is it warm?

Rules

$\forall d \; Sunny(d) \Rightarrow Warm(d)$

Wrong; it can be sunny but cold

$\forall d \; Warm(d) \Rightarrow Sunny(d)$

Wrong; it can be warm and cloudy

Problem

Can't capture that sunny days are usually warm

Probability Theory

Key Idea

Measure degrees of belief in propositions

Random Variable	An interesting part of the world e.g. *Sky* or *Temp*
Domain	Possible values of a random variable e.g. $domain(Temp) = \langle warm, cold, \ldots \rangle$
Proposition	A statement, like propositional logic e.g. $Sky = sunny \wedge Temp = warm$
Probability	The degree of belief in a proposition e.g. $P(Temp = warm) = 0.7$

Probability Theory

Key Idea

Measure degrees of belief in propositions

Random Variable	An interesting part of the world e.g. *Sky* or *Temp*
Domain	Possible values of a random variable e.g. $domain(Temp) = \langle warm, cold, \ldots \rangle$
Proposition	A statement, like propositional logic e.g. $Sky = sunny \wedge Temp = warm$
Probability	The degree of belief in a proposition e.g. $P(Temp = warm) = 0.7$

Probability Theory

Key Idea

Measure degrees of belief in propositions

Random Variable	An interesting part of the world e.g. *Sky* or *Temp*
Domain	Possible values of a random variable e.g. $domain(Temp) = \langle warm, cold, \ldots \rangle$
Proposition	A statement, like propositional logic e.g. $Sky = sunny \wedge Temp = warm$
Probability	The degree of belief in a proposition e.g. $P(Temp = warm) = 0.7$

Probability Theory

Key Idea

Measure degrees of belief in propositions

Random Variable An interesting part of the world
e.g. *Sky* or *Temp*

Domain Possible values of a random variable
e.g. $domain(Temp) = \langle warm, cold, \ldots \rangle$

Proposition A statement, like propositional logic
e.g. $Sky = sunny \wedge Temp = warm$

Probability The degree of belief in a proposition
e.g. $P(Temp = warm) = 0.7$

Probability Theory

Key Idea

Measure degrees of belief in propositions

Random Variable	An interesting part of the world e.g. *Sky* or *Temp*
Domain	Possible values of a random variable e.g. $domain(Temp) = \langle warm, cold, \ldots \rangle$
Proposition	A statement, like propositional logic e.g. $Sky = sunny \wedge Temp = warm$
Probability	The degree of belief in a proposition e.g. $P(Temp = warm) = 0.7$

Aside: Fuzzy Logic

Probability Theory

- Propositions are believed to a certain degree
- Belief values range from [0, 1]

Fuzzy Logic

- Propositions are *true* to a certain degree
- *Truth* values range from [0, 1]

For example:

- $T(Tall(Steve)) = 0.5$
- $T(Fat(Steve)) = 0.1$

Better for describing indefinite classes than for reasoning

Aside: Fuzzy Logic

Probability Theory

- Propositions are believed to a certain degree
- Belief values range from [0, 1]

Fuzzy Logic

- Propositions are *true* to a certain degree
- *Truth* values range from [0, 1]

For example:

- $T(Tall(Steve)) = 0.5$
- $T(Fat(Steve)) = 0.1$

Better for describing indefinite classes than for reasoning

Aside: Fuzzy Logic

Probability Theory

- Propositions are believed to a certain degree
- Belief values range from [0, 1]

Fuzzy Logic

- Propositions are *true* to a certain degree
- *Truth* values range from [0, 1]

For example:

- $T(\mathit{Tall}(\mathit{Steve})) = 0.5$
- $T(\mathit{Fat}(\mathit{Steve})) = 0.1$

Better for describing indefinite classes than for reasoning

Aside: Fuzzy Logic

Probability Theory

- Propositions are believed to a certain degree
- Belief values range from [0, 1]

Fuzzy Logic

- Propositions are *true* to a certain degree
- *Truth* values range from [0, 1]

For example:

- $T(\mathit{Tall}(\mathit{Steve})) = 0.5$
- $T(\mathit{Fat}(\mathit{Steve})) = 0.1$

Better for describing indefinite classes than for reasoning

Probability Rules

Range of Probabilities

$0 \leq P(a) \leq 1$

Propositions Known to be True or False

$$\begin{aligned} P(true) &= 1 \\ P(false) &= 0 \end{aligned}$$

Probability of Disjunctions

$P(a \vee b) = P(a) + P(b) - P(a \wedge b)$

Probability Rules

Range of Probabilities

$0 \leq P(a) \leq 1$

Propositions Known to be True or False

$$\begin{aligned} P(\mathit{true}) &= 1 \\ P(\mathit{false}) &= 0 \end{aligned}$$

Probability of Disjunctions

$P(a \vee b) = P(a) + P(b) - P(a \wedge b)$

Probability Rules

Range of Probabilities

$$0 \leq P(a) \leq 1$$

Propositions Known to be True or False

$$\begin{aligned} P(\mathit{true}) &= 1 \\ P(\mathit{false}) &= 0 \end{aligned}$$

Probability of Disjunctions

$$P(a \vee b) = P(a) + P(b) - P(a \wedge b)$$

Probability Rules

Range of Probabilities

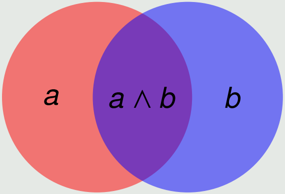
$0 \leq P(a) \leq 1$

Propositions Known to be True or False

$$\begin{aligned} P(\mathit{true}) &= 1 \\ P(\mathit{false}) &= 0 \end{aligned}$$

Probability of Disjunctions

$P(a \vee b) = P(a) + P(b) - P(a \wedge b)$



Random Variables

Boolean Random Variables

Have the domain $\langle true, false \rangle$, e.g.

- *IsSunny*

Discrete Random Variables

Have a countable domain, e.g.

- $domain(Sky) = \langle sunny, cloudy, \ldots \rangle$
- $domain(DieRoll) = \langle 1, 2, 3, 4, 5, 6 \rangle$

Continuous Random Variables

Have a real-valued domain, e.g.

- $domain(Temperature) = \langle \ldots, -40^\circ, \ldots, 98.6^\circ, \ldots \rangle$

Random Variables

Boolean Random Variables

Have the domain $\langle true, false \rangle$, e.g.

- $IsSunny$

Discrete Random Variables

Have a countable domain, e.g.

- $domain(Sky) = \langle sunny, cloudy, \ldots \rangle$
- $domain(DieRoll) = \langle 1, 2, 3, 4, 5, 6 \rangle$

Continuous Random Variables

Have a real-valued domain, e.g.

- $domain(Temperature) = \langle \ldots, -40^\circ, \ldots, 98.6^\circ, \ldots \rangle$

Random Variables

Boolean Random Variables

Have the domain $\langle \mathit{true}, \mathit{false} \rangle$, e.g.

- *IsSunny*

Discrete Random Variables

Have a countable domain, e.g.

- $\mathit{domain}(\mathit{Sky}) = \langle \mathit{sunny}, \mathit{cloudy}, \ldots \rangle$
- $\mathit{domain}(\mathit{DieRoll}) = \langle 1, 2, 3, 4, 5, 6 \rangle$

Continuous Random Variables

Have a real-valued domain, e.g.

- $\mathit{domain}(\mathit{Temperature}) = \langle \ldots, -40^\circ, \ldots, 98.6^\circ, \ldots \rangle$

Prior Probability

Definition

$P(a)$ The unconditional or prior probability of a proposition a is the degree of belief in that proposition *given no other information*

Prior Probability

Definition

$P(a)$ The unconditional or prior probability of a proposition a is the degree of belief in that proposition *given no other information*

Examples

$P(DieRoll = 5) \quad =$

Prior Probability

Definition

$P(a)$ The unconditional or prior probability of a proposition a is the degree of belief in that proposition *given no other information*

Examples

$$P(DieRoll = 5) = 1/6$$

Prior Probability

Definition

$P(a)$ The unconditional or prior probability of a proposition a is the degree of belief in that proposition *given no other information*

Examples

$$
\begin{aligned}
P(\mathit{DieRoll} = 5) &= 1/6 \\
P(\mathit{CardDrawn} = \mathrm{A}\spadesuit) &=
\end{aligned}
$$

Prior Probability

Definition

$P(a)$ The unconditional or prior probability of a proposition a is the degree of belief in that proposition *given no other information*

Examples

$$P(DieRoll = 5) = 1/6$$
$$P(CardDrawn = A\spadesuit) = 1/52$$

Prior Probability

Definition

$P(a)$ The unconditional or prior probability of a proposition a is the degree of belief in that proposition *given no other information*

Examples

$$
\begin{aligned}
P(\mathit{DieRoll} = 5) &= 1/6 \\
P(\mathit{CardDrawn} = \mathrm{A}\spadesuit) &= 1/52 \\
P(\mathit{SkyInBirmingham} = \mathit{sunny}) &=
\end{aligned}
$$

Prior Probability

Definition

$P(a)$ The unconditional or prior probability of a proposition a is the degree of belief in that proposition *given no other information*

Examples

$$
\begin{aligned}
P(\mathit{DieRoll} = 5) &= 1/6 \\
P(\mathit{CardDrawn} = \mathrm{A}\spadesuit) &= 1/52 \\
P(\mathit{SkyInBirmingham} = \mathit{sunny}) &= 210/365
\end{aligned}
$$

Joint Probability

Definition

$P(a_1, \ldots, a_n)$ The joint probability of propositions $a_1, \ldots, a_n$ is the degree of belief in the proposition $a_1 \wedge \ldots \wedge a_n$

Joint Probability

Definition

$P(a_1, \ldots, a_n)$ The **joint probability** of propositions $a_1, \ldots, a_n$ is the degree of belief in the proposition $a_1 \wedge \ldots \wedge a_n$

Example

$$
\begin{aligned}
& P(\mathit{DieRoll}_1 = 4, \mathit{DieRoll}_2 = 6) \\
& \quad = \quad P(\mathit{DieRoll}_1 = 4 \wedge \mathit{DieRoll}_2 = 6) \\
& \quad =
\end{aligned}
$$

Joint Probability

Definition

$P(a_1, \ldots, a_n)$ The joint probability of propositions $a_1, \ldots, a_n$ is the degree of belief in the proposition $a_1 \wedge \ldots \wedge a_n$

Example

$$
\begin{aligned}
& P(\mathit{DieRoll}_1 = 4, \mathit{DieRoll}_2 = 6) \\
& \quad = \quad P(\mathit{DieRoll}_1 = 4 \wedge \mathit{DieRoll}_2 = 6) \\
& \quad = \quad 1/36
\end{aligned}
$$

Conditional Probability

Definition

$P(a|b)$ The posterior or conditional probability of a proposition a given a proposition b is the degree of belief in a, *given that we know only* b

Conditional Probability

Definition

$P(a|b)$ The posterior or conditional probability of a proposition a given a proposition b is the degree of belief in a, *given that we know only* b

Examples

$$P(Card = \text{A}\spadesuit | CardSuit = \spadesuit) \quad =$$

Conditional Probability

Definition

$P(a|b)$ The posterior or conditional probability of a proposition a given a proposition b is the degree of belief in a, *given that we know only* b

Examples

$$P(Card = \text{A}\spadesuit | CardSuit = \spadesuit) = 1/13$$

Conditional Probability

Definition

$P(a|b)$ The posterior or conditional probability of a proposition a given a proposition b is the degree of belief in a, *given that we know only* b

Examples

$$
\begin{aligned}
P(Card = \mathrm{A}\spadesuit | CardSuit = \spadesuit) &= 1/13 \\
P(DieRoll_2 = 5 | DieRoll_1 = 5) &=
\end{aligned}
$$

Conditional Probability

Definition

$P(a|b)$ The posterior or conditional probability of a proposition a given a proposition b is the degree of belief in a, *given that we know only* b

Examples

$$
\begin{aligned}
P(\mathit{Card} = \text{A}\spadesuit | \mathit{CardSuit} = \spadesuit) &= 1/13 \\
P(\mathit{DieRoll}_2 = 5 | \mathit{DieRoll}_1 = 5) &= 1/6
\end{aligned}
$$

Prior, Joint or Conditional?

1. Probability of a having a cavity?
   Prior, $P(Cavity = true)$
2. Probability of it being warm and cloudy?
   Joint, $P(Temp = warm, Sky = cloudy)$
3. Probability of car being stolen?
   Prior, $P(CarStolen = true)$
4. Probability of car being stolen and being in Alabama?
   Joint, $P(CarStolen = true, InAlabama = true)$
5. The car is in Alabama. Probability of car being stolen?
   Conditional, $P(CarStolen = true | InAlabama = true)$
6. It's cloudy. Probability of it being warm?
   Conditional, $P(Temp = warm | Sky = cloudy)$

Prior, Joint or Conditional?

1. Probability of a having a cavity?
   Prior, $P(Cavity = true)$
2. Probability of it being warm and cloudy?
   Joint, $P(Temp = warm, Sky = cloudy)$
3. Probability of car being stolen?
   Prior, $P(CarStolen = true)$
4. Probability of car being stolen and being in Alabama?
   Joint, $P(CarStolen = true, InAlabama = true)$
5. The car is in Alabama. Probability of car being stolen?
   Conditional, $P(CarStolen = true | InAlabama = true)$
6. It's cloudy. Probability of it being warm?
   Conditional, $P(Temp = warm | Sky = cloudy)$

Prior, Joint or Conditional?

1. Probability of a having a cavity?
   Prior, $P(Cavity = true)$
2. Probability of it being warm and cloudy?
   Joint, $P(Temp = warm, Sky = cloudy)$
3. Probability of car being stolen?
   Prior, $P(CarStolen = true)$
4. Probability of car being stolen and being in Alabama?
   Joint, $P(CarStolen = true, InAlabama = true)$
5. The car is in Alabama. Probability of car being stolen?
   Conditional, $P(CarStolen = true | InAlabama = true)$
6. It's cloudy. Probability of it being warm?
   Conditional, $P(Temp = warm | Sky = cloudy)$

Prior, Joint or Conditional?

1. Probability of a having a cavity?
   Prior, $P(Cavity = true)$
2. Probability of it being warm and cloudy?
   Joint, $P(Temp = warm, Sky = cloudy)$
3. Probability of car being stolen?
   Prior, $P(CarStolen = true)$
4. Probability of car being stolen and being in Alabama?
   Joint, $P(CarStolen = true, InAlabama = true)$
5. The car is in Alabama. Probability of car being stolen?
   Conditional, $P(CarStolen = true | InAlabama = true)$
6. It's cloudy. Probability of it being warm?
   Conditional, $P(Temp = warm | Sky = cloudy)$

Prior, Joint or Conditional?

1. Probability of a having a cavity?
   Prior, $P(Cavity = true)$
2. Probability of it being warm and cloudy?
   Joint, $P(Temp = warm, Sky = cloudy)$
3. Probability of car being stolen?
   Prior, $P(CarStolen = true)$
4. Probability of car being stolen and being in Alabama?
   Joint, $P(CarStolen = true, InAlabama = true)$
5. The car is in Alabama. Probability of car being stolen?
   Conditional, $P(CarStolen = true | InAlabama = true)$
6. It's cloudy. Probability of it being warm?
   Conditional, $P(Temp = warm | Sky = cloudy)$

Prior, Joint or Conditional?

1. Probability of a having a cavity?
   Prior, $P(Cavity = true)$
2. Probability of it being warm and cloudy?
   Joint, $P(Temp = warm, Sky = cloudy)$
3. Probability of car being stolen?
   Prior, $P(CarStolen = true)$
4. Probability of car being stolen and being in Alabama?
   Joint, $P(CarStolen = true, InAlabama = true)$
5. The car is in Alabama. Probability of car being stolen?
   Conditional, $P(CarStolen = true | InAlabama = true)$
6. It's cloudy. Probability of it being warm?
   Conditional, $P(Temp = warm | Sky = cloudy)$

Prior, Joint or Conditional?

1. Probability of a having a cavity?
   Prior, $P(Cavity = true)$
2. Probability of it being warm and cloudy?
   Joint, $P(Temp = warm, Sky = cloudy)$
3. Probability of car being stolen?
   Prior, $P(CarStolen = true)$
4. Probability of car being stolen and being in Alabama?
   Joint, $P(CarStolen = true, InAlabama = true)$
5. The car is in Alabama. Probability of car being stolen?
   Conditional, $P(CarStolen = true | InAlabama = true)$
6. It's cloudy. Probability of it being warm?
   Conditional, $P(Temp = warm | Sky = cloudy)$

Prior, Joint or Conditional?

1. Probability of a having a cavity?
   Prior, $P(Cavity = true)$
2. Probability of it being warm and cloudy?
   Joint, $P(Temp = warm, Sky = cloudy)$
3. Probability of car being stolen?
   Prior, $P(CarStolen = true)$
4. Probability of car being stolen and being in Alabama?
   Joint, $P(CarStolen = true, InAlabama = true)$
5. The car is in Alabama. Probability of car being stolen?
   Conditional, $P(CarStolen = true | InAlabama = true)$
6. It's cloudy. Probability of it being warm?
   Conditional, $P(Temp = warm | Sky = cloudy)$

Prior, Joint or Conditional?

1. Probability of a having a cavity?
   Prior, $P(Cavity = true)$
2. Probability of it being warm and cloudy?
   Joint, $P(Temp = warm, Sky = cloudy)$
3. Probability of car being stolen?
   Prior, $P(CarStolen = true)$
4. Probability of car being stolen and being in Alabama?
   Joint, $P(CarStolen = true, InAlabama = true)$
5. The car is in Alabama. Probability of car being stolen?
   Conditional, $P(CarStolen = true | InAlabama = true)$
6. It's cloudy. Probability of it being warm?
   Conditional, $P(Temp = warm | Sky = cloudy)$

Prior, Joint or Conditional?

1. Probability of a having a cavity?
   Prior, $P(Cavity = true)$
2. Probability of it being warm and cloudy?
   Joint, $P(Temp = warm, Sky = cloudy)$
3. Probability of car being stolen?
   Prior, $P(CarStolen = true)$
4. Probability of car being stolen and being in Alabama?
   Joint, $P(CarStolen = true, InAlabama = true)$
5. The car is in Alabama. Probability of car being stolen?
   Conditional, $P(CarStolen = true | InAlabama = true)$
6. It's cloudy. Probability of it being warm?
   Conditional, $P(Temp = warm | Sky = cloudy)$

Prior, Joint or Conditional?

1. Probability of a having a cavity?
   Prior, $P(Cavity = true)$
2. Probability of it being warm and cloudy?
   Joint, $P(Temp = warm, Sky = cloudy)$
3. Probability of car being stolen?
   Prior, $P(CarStolen = true)$
4. Probability of car being stolen and being in Alabama?
   Joint, $P(CarStolen = true, InAlabama = true)$
5. The car is in Alabama. Probability of car being stolen?
   Conditional, $P(CarStolen = true | InAlabama = true)$
6. It's cloudy. Probability of it being warm?
   Conditional, $P(Temp = warm | Sky = cloudy)$

Prior, Joint or Conditional?

1. Probability of a having a cavity?
   Prior, $P(Cavity = true)$
2. Probability of it being warm and cloudy?
   Joint, $P(Temp = warm, Sky = cloudy)$
3. Probability of car being stolen?
   Prior, $P(CarStolen = true)$
4. Probability of car being stolen and being in Alabama?
   Joint, $P(CarStolen = true, InAlabama = true)$
5. The car is in Alabama. Probability of car being stolen?
   Conditional, $P(CarStolen = true | InAlabama = true)$
6. It's cloudy. Probability of it being warm?
   Conditional, $P(Temp = warm | Sky = cloudy)$

Relation between Joint and Conditional

Product Rule

$P(a \wedge b) = P(a|b)P(b)$ or $P(a|b) = P(a \wedge b)/P(b)$

Relation between Joint and Conditional

Product Rule

$P(a \wedge b) = P(a|b)P(b)$ or $P(a|b) = P(a \wedge b)/P(b)$

Intuition

To have $a \wedge b$ true, we need b true, and a true given b

Relation between Joint and Conditional

Product Rule

$P(a \wedge b) = P(a|b)P(b)$ or $P(a|b) = P(a \wedge b)/P(b)$

Intuition

To have $a \wedge b$ true, we need b true, and a true given b

Example

$$P(\mathrm{A} \wedge \spadesuit) \quad =$$

Relation between Joint and Conditional

Product Rule

$P(a \wedge b) = P(a|b)P(b)$ or $P(a|b) = P(a \wedge b)/P(b)$

Intuition

To have $a \wedge b$ true, we need b true, and a true given b

Example

$$P(\mathrm{A} \wedge \spadesuit) = 1/52$$

Relation between Joint and Conditional

Product Rule

$P(a \wedge b) = P(a|b)P(b)$ or $P(a|b) = P(a \wedge b)/P(b)$

Intuition

To have $a \wedge b$ true, we need b true, and a true given b

Example

$$
\begin{aligned}
P(\mathrm{A} \wedge \spadesuit) &= 1/52 \\
P(\mathrm{A}|\spadesuit) &=
\end{aligned}
$$

Relation between Joint and Conditional

Product Rule

$P(a \wedge b) = P(a|b)P(b)$ or $P(a|b) = P(a \wedge b)/P(b)$

Intuition

To have $a \wedge b$ true, we need b true, and a true given b

Example

$$
\begin{aligned}
P(\mathrm{A} \wedge \spadesuit) &= 1/52 \\
P(\mathrm{A}|\spadesuit) &= 1/13
\end{aligned}
$$

Relation between Joint and Conditional

Product Rule

$P(a \wedge b) = P(a|b)P(b)$ or $P(a|b) = P(a \wedge b)/P(b)$

Intuition

To have $a \wedge b$ true, we need b true, and a true given b

Example

$$
\begin{aligned}
P(\mathrm{A} \wedge \spadesuit) &= 1/52 \\
P(\mathrm{A}|\spadesuit) &= 1/13 \\
P(\spadesuit) &=
\end{aligned}
$$

Relation between Joint and Conditional

Product Rule

$P(a \wedge b) = P(a|b)P(b)$ or $P(a|b) = P(a \wedge b)/P(b)$

Intuition

To have $a \wedge b$ true, we need b true, and a true given b

Example

$$
\begin{aligned}
P(\mathrm{A} \wedge \spadesuit) &= 1/52 \\
P(\mathrm{A}|\spadesuit) &= 1/13 \\
P(\spadesuit) &= 1/4
\end{aligned}
$$

Relation between Joint and Conditional

Product Rule

$P(a \wedge b) = P(a|b)P(b)$ or $P(a|b) = P(a \wedge b)/P(b)$

Intuition

To have $a \wedge b$ true, we need b true, and a true given b

Example

$$
\begin{aligned}
P(\mathrm{A} \wedge \spadesuit) &= 1/52 \\
P(\mathrm{A}|\spadesuit) &= 1/13 \\
P(\spadesuit) &= 1/4 \\
P(\mathrm{A}|\spadesuit)P(\spadesuit) &=
\end{aligned}
$$

Relation between Joint and Conditional

Product Rule

$P(a \wedge b) = P(a|b)P(b)$ or $P(a|b) = P(a \wedge b)/P(b)$

Intuition

To have $a \wedge b$ true, we need b true, and a true given b

Example

$$
\begin{aligned}
P(\mathrm{A} \wedge \spadesuit) &= 1/52 \\
P(\mathrm{A}|\spadesuit) &= 1/13 \\
P(\spadesuit) &= 1/4 \\
P(\mathrm{A}|\spadesuit)P(\spadesuit) &= 1/13 \cdot 1/4 = 1/52
\end{aligned}
$$

Chain Rule

Key Ideas

- Repeatedly apply product rule, $P(a, b) = P(a|b)P(b)$
- Joint probability $\rightarrow$ conditional probabilities

Chain Rule

Key Ideas

- Repeatedly apply product rule, $P(a, b) = P(a|b)P(b)$
- Joint probability $\rightarrow$ conditional probabilities

Example

$P(sunny, dry, warm)$

Chain Rule

Key Ideas

- Repeatedly apply product rule, $P(a, b) = P(a|b)P(b)$
- Joint probability $\rightarrow$ conditional probabilities

Example

$$P(sunny, dry, warm) = P(sunny|dry, warm)P(dry, warm)$$

Chain Rule

Key Ideas

- Repeatedly apply product rule, $P(a, b) = P(a|b)P(b)$
- Joint probability $\rightarrow$ conditional probabilities

Example

$$\begin{aligned} & P(sunny, dry, warm) \\ &= P(sunny|dry, warm)P(dry, warm) \\ &= P(sunny|dry, warm)P(dry|warm)P(warm) \end{aligned}$$

Outline

Probability Distributions

Definition

$\mathbf{P}(X)$ The probability distribution of a random variable X is a list of probabilities for each domain value

Probability Distributions

Definition

$\mathbf{P}(X)$ The probability distribution of a random variable X is a list of probabilities for each domain value

Example

$\mathbf{P}(\mathit{SkyInBirmingham}) = \langle 210/365, 155/365 \rangle$ means

$$
\begin{aligned}
P(\mathit{SkyInBirmingham} = \mathit{sunny}) &= 210/365 \\
P(\mathit{SkyInBirmingham} = \mathit{cloudy}) &= 155/365
\end{aligned}
$$

Probability Distributions

Definition

$\mathbf{P}(X)$ The probability distribution of a random variable X is a list of probabilities for each domain value

Example

$\mathbf{P}(\mathit{SkyInBirmingham}) = \langle 210/365, 155/365 \rangle$ means

$$\begin{aligned} P(\mathit{SkyInBirmingham} = \mathit{sunny}) &= 210/365 \\ P(\mathit{SkyInBirmingham} = \mathit{cloudy}) &= 155/365 \end{aligned}$$

Notation Warning

- $P(a)$ or $P(X = a)$ means prior probability
- $\mathbf{P}(X)$ means probability distribution

Probability Distributions

Key Idea

$$\sum_{i}^{d} P(X = X_i) = 1$$

The sum of the probabilities for all possible value assignments of the random variable is always 1

Probability Distributions

Key Idea

$$\sum_i^d P(X = X_i) = 1$$

The sum of the probabilities for all possible value assignments of the random variable is always 1

Example

$$\sum_x P(Die = x)$$

Probability Distributions

Key Idea

$$\sum_{i}^{d} P(X = X_i) = 1$$

The sum of the probabilities for all possible value assignments of the random variable is always 1

Example

$$\begin{aligned}\sum_{x} P(Die = x) \\ = \quad & P(Die = 1) + P(Die = 2) + P(Die = 3) + \\ & P(Die = 4) + P(Die = 5) + P(Die = 6)\end{aligned}$$

Probability Distributions

Key Idea

$$\sum_{i}^{d} P(X = X_i) = 1$$

The sum of the probabilities for all possible value assignments of the random variable is always 1

Example

$$\begin{aligned}
\sum_{x} P(Die = x) \\
&= P(Die = 1) + P(Die = 2) + P(Die = 3) + \\
&\quad P(Die = 4) + P(Die = 5) + P(Die = 6) \\
&= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6}
\end{aligned}$$

Probability Distributions

Key Idea

$$\sum_{i}^{d} P(X = X_i) = 1$$

The sum of the probabilities for all possible value assignments of the random variable is always 1

Example

$$\begin{aligned}
\sum_{x} P(Die = x) \\
&= P(Die = 1) + P(Die = 2) + P(Die = 3) + \\
&\quad P(Die = 4) + P(Die = 5) + P(Die = 6) \\
&= \tfrac{1}{6} + \tfrac{1}{6} + \tfrac{1}{6} + \tfrac{1}{6} + \tfrac{1}{6} + \tfrac{1}{6} \\
&= 1
\end{aligned}$$

Continuous Variable Probability Distributions

Definition

A probability density function is a probability distribution over a continuous variable

Key Idea

Function assigns a probability to all possible values

Example

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Continuous Variable Probability Distributions

Definition

A probability density function is a probability distribution over a continuous variable

Key Idea

Function assigns a probability to all possible values

Example

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Continuous Variable Probability Distributions

Definition

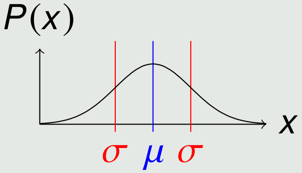
A probability density function is a probability distribution over a continuous variable

Key Idea

Function assigns a probability to all possible values

Example

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



Joint Probability Distributions

Definition

$\mathbf{P}(X_1, \ldots, X_n)$

The joint probability distribution of random variables $X_1, \ldots, X_n$ is a table of probabilities for each combination of values in the variable domains

Example

$\mathbf{P}(Sex, Smoker) =$

	$Sex = male$	$Sex = female$
$Smoker = true$	0.113	0.107
$Smoker = false$	0.377	0.403

Joint Probability Distributions

Definition

$\mathbf{P}(X_1, \ldots, X_n)$

The joint probability distribution of random variables $X_1, \ldots, X_n$ is a table of probabilities for each combination of values in the variable domains

Example

$\mathbf{P}(Sex, Smoker) =$

	$Sex = male$	$Sex = female$
$Smoker = true$	0.113	0.107
$Smoker = false$	0.377	0.403

Inference with Joint Probability Distributions

Key Idea

Sum entries in joint distribution where proposition is true

Inference with Joint Probability Distributions

Key Idea

Sum entries in joint distribution where proposition is true

Example: $P(Sex = female \lor Smoker = false)$

	$Sex = male$	$Sex = female$
$Smoker = true$	0.113	0.107
$Smoker = false$	0.377	0.403

Inference with Joint Probability Distributions

Key Idea

Sum entries in joint distribution where proposition is true

Example: $P(Sex = female \vee Smoker = false)$

	$Sex = male$	$Sex = female$
$Smoker = true$	0.113	0.107
$Smoker = false$	0.377	0.403

$P(Sex = female \vee Smoker = false)$

Inference with Joint Probability Distributions

Key Idea

Sum entries in joint distribution where proposition is true

Example: $P(Sex = female \vee Smoker = false)$

	$Sex = male$	$Sex = female$
$Smoker = true$	0.113	0.107
$Smoker = false$	0.377	0.403

$$\begin{aligned} & P(Sex = female \vee Smoker = false) \\ & \quad = \; P(Sex = male, Smoker = false) + \\ & \qquad\; P(Sex = female, Smoker = true) + \\ & \qquad\; P(Sex = female, Smoker = false) \end{aligned}$$

Inference with Joint Probability Distributions

Key Idea

Sum entries in joint distribution where proposition is true

Example: $P(Sex = female \vee Smoker = false)$

	$Sex = male$	$Sex = female$
$Smoker = true$	0.113	0.107
$Smoker = false$	0.377	0.403

$$\begin{aligned} & P(Sex = female \vee Smoker = false) \\ = \; & P(Sex = male, Smoker = false) + \\ & P(Sex = female, Smoker = true) + \\ & P(Sex = female, Smoker = false) \\ = \; & 0.377 + 0.107 + 0.403 \end{aligned}$$

Inference with Joint Probability Distributions

Key Idea

Sum entries in joint distribution where proposition is true

Example: $P(Sex = female \vee Smoker = false)$

	$Sex = male$	$Sex = female$
$Smoker = true$	0.113	0.107
$Smoker = false$	0.377	0.403

$$
\begin{aligned}
& P(Sex = female \vee Smoker = false) \\
= \;& P(Sex = male, Smoker = false) + \\
& P(Sex = female, Smoker = true) + \\
& P(Sex = female, Smoker = false) \\
= \;& 0.377 + 0.107 + 0.403 = 0.887
\end{aligned}
$$

Marginalization

Key Idea

$$\mathbf{P}(Y) = \sum_z \mathbf{P}(Y, z)$$

Marginalization removes all variables but $\mathbf{Y}$ by summing over the values of the other variables

Marginalization

Key Idea

$$\mathbf{P}(Y) = \sum_z \mathbf{P}(Y, z)$$

Marginalization removes all variables but **Y** by summing over the values of the other variables

Example: $P(Sex = female)$

The other variable is *Smoker*, so:

$P(Sex = female)$

Marginalization

Key Idea

$$\mathbf{P}(Y) = \sum_z \mathbf{P}(Y, z)$$

Marginalization removes all variables but $\mathbf{Y}$ by summing over the values of the other variables

Example: $P(Sex = female)$

The other variable is *Smoker*, so:

$$\begin{aligned} &P(Sex = female) \\ &\quad = \; P(Sex = female, Smoker = true) + \\ &\qquad\;\; P(Sex = female, Smoker = false) \end{aligned}$$

Marginalization

Key Idea

$$\mathbf{P}(Y) = \sum_z \mathbf{P}(Y, z)$$

Marginalization removes all variables but **Y** by summing over the values of the other variables

Example: $P(Sex = female)$

The other variable is *Smoker*, so:

$$\begin{aligned} & P(Sex = female) \\ = \quad & P(Sex = female, Smoker = true) + \\ & P(Sex = female, Smoker = false) \\ = \quad & 0.107 + 0.403 \end{aligned}$$

Marginalization

Key Idea

$$\mathbf{P}(Y) = \sum_z \mathbf{P}(Y, z)$$

Marginalization removes all variables but $\mathbf{Y}$ by summing over the values of the other variables

Example: $P(Sex = female)$

The other variable is *Smoker*, so:

$$\begin{aligned} P(Sex = female) & \\ = \; & P(Sex = female, Smoker = true) + \\ & P(Sex = female, Smoker = false) \\ = \; & 0.107 + 0.403 = 0.51 \end{aligned}$$

Normalization

Key Idea

$$\mathbf{P}(Y|z) = \frac{\mathbf{P}(Y,z)}{P(z)} = \frac{\mathbf{P}(Y,z)}{\sum_y P(y,z)} = \alpha\mathbf{P}(Y,z)$$

Normalization

Key Idea

$$\mathbf{P}(Y|z) = \frac{\mathbf{P}(Y,z)}{P(z)} = \frac{\mathbf{P}(Y,z)}{\sum_y P(y,z)} = \alpha\mathbf{P}(Y,z)$$

Calculating a Normalizing Constant

$$\begin{aligned} P(\spadesuit, \mathrm{A}) &= \frac{1}{52} \\ P(\clubsuit, \mathrm{A}) &= \frac{1}{52} \\ P(\diamondsuit, \mathrm{A}) &= \frac{1}{52} \\ P(\heartsuit, \mathrm{A}) &= \frac{1}{52} \end{aligned}$$

Normalization

Key Idea

$$\mathbf{P}(Y|z) = \frac{\mathbf{P}(Y,z)}{P(z)} = \frac{\mathbf{P}(Y,z)}{\sum_y P(y,z)} = \alpha \mathbf{P}(Y,z)$$

Calculating a Normalizing Constant

$$\begin{aligned} P(\spadesuit, \mathrm{A}) &= \frac{1}{52} \\ P(\clubsuit, \mathrm{A}) &= \frac{1}{52} \\ P(\diamondsuit, \mathrm{A}) &= \frac{1}{52} \\ P(\heartsuit, \mathrm{A}) &= \frac{1}{52} \end{aligned}$$

$$\alpha =$$

Normalization

Key Idea

$$\mathbf{P}(Y|z) = \frac{\mathbf{P}(Y,z)}{P(z)} = \frac{\mathbf{P}(Y,z)}{\sum_y P(y,z)} = \alpha \mathbf{P}(Y,z)$$

Calculating a Normalizing Constant

$$\begin{aligned} P(\spadesuit, \mathrm{A}) &= \frac{1}{52} \\ P(\clubsuit, \mathrm{A}) &= \frac{1}{52} \\ P(\diamondsuit, \mathrm{A}) &= \frac{1}{52} \\ P(\heartsuit, \mathrm{A}) &= \frac{1}{52} \end{aligned}$$

$$\alpha = \frac{1}{\frac{1}{52} + \frac{1}{52} + \frac{1}{52} + \frac{1}{52}} = 13$$

Normalization

Key Idea

$$\mathbf{P}(Y|z) = \frac{\mathbf{P}(Y,z)}{P(z)} = \frac{\mathbf{P}(Y,z)}{\sum_y P(y,z)} = \alpha \mathbf{P}(Y,z)$$

Calculating a Normalizing Constant

$$\begin{aligned} P(\spadesuit, \mathrm{A}) &= \frac{1}{52} \\ P(\clubsuit, \mathrm{A}) &= \frac{1}{52} \\ P(\diamondsuit, \mathrm{A}) &= \frac{1}{52} \\ P(\heartsuit, \mathrm{A}) &= \frac{1}{52} \end{aligned}$$

$$\alpha = \frac{1}{\frac{1}{52} + \frac{1}{52} + \frac{1}{52} + \frac{1}{52}} = 13$$

$$\mathbf{P}(Suit|\mathrm{A}) = \langle \frac{13}{52}, \frac{13}{52}, \frac{13}{52}, \frac{13}{52} \rangle = \langle \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4} \rangle$$

Inference Exercises

$sunny$	$warm$	$snow$	0.01
$sunny$	$warm$	$\neg snow$	0.59
$sunny$	$cold$	$snow$	0.08
$sunny$	$cold$	$\neg snow$	0.14
$cloudy$	$warm$	$snow$	0.03
$cloudy$	$warm$	$\neg snow$	0.07
$cloudy$	$cold$	$snow$	0.06
$cloudy$	$cold$	$\neg snow$	0.02

1. $P(sunny) =$
2. $P(warm) =$
3. $P(snow) =$
4. $P(sunny \vee \neg snow) =$
5. $P(sunny | snow) =$
6. $P(snow | sunny, cold) =$

Inference Exercises

$sunny$	$warm$	$snow$	0.01
$sunny$	$warm$	$\neg snow$	0.59
$sunny$	$cold$	$snow$	0.08
$sunny$	$cold$	$\neg snow$	0.14
$cloudy$	$warm$	$snow$	0.03
$cloudy$	$warm$	$\neg snow$	0.07
$cloudy$	$cold$	$snow$	0.06
$cloudy$	$cold$	$\neg snow$	0.02

1. $P(sunny) = 0.82$
2. $P(warm) = 0.70$
3. $P(snow) = 0.18$
4. $P(sunny \vee \neg snow) = 0.91$
5. $P(sunny|snow) = 0.50$
6. $P(snow|sunny, cold) = 0.36$

Joint Distribution Inference

Properties

Given n random variables with maximum domain size d:

Time Complexity? $O(d^n)$

Space Complexity? $O(d^n)$

Biggest Problem

How do you fill in a table of $O(d^n)$ probabilities?

Joint Distribution Inference

Properties

Given n random variables with maximum domain size d:

Time Complexity? $O(d^n)$

Space Complexity? $O(d^n)$

Biggest Problem

How do you fill in a table of $O(d^n)$ probabilities?

Joint Distribution Inference

Properties

Given n random variables with maximum domain size d:

Time Complexity? $O(d^n)$

Space Complexity? $O(d^n)$

Biggest Problem

How do you fill in a table of $O(d^n)$ probabilities?

Joint Distribution Inference

Properties

Given n random variables with maximum domain size d:

Time Complexity? $O(d^n)$

Space Complexity? $O(d^n)$

Biggest Problem

How do you fill in a table of $O(d^n)$ probabilities?

Outline

Independence

Key Ideas

- Sometimes no connection exists between variables
- Such independence determined by world knowledge

Formal Independence

$P(a, b) = P(a)P(b)$ or $P(a|b) = P(a)$

Examples

$$
\begin{aligned}
P(EyeColor, Gender) &= P(EyeColor)P(Gender) \\
P(Cavity, BlazersWon) &= P(Cavity)P(BlazersWon) \\
P(DieRoll_1, DieRoll_2) &= P(DieRoll_1)P(DieRoll_2)
\end{aligned}
$$

Independence

Key Ideas

- Sometimes no connection exists between variables
- Such independence determined by world knowledge

Formal Independence

$P(a, b) = P(a)P(b)$ or $P(a|b) = P(a)$

Examples

$$
\begin{aligned}
P(EyeColor, Gender) &= P(EyeColor)P(Gender) \\
P(Cavity, BlazersWon) &= P(Cavity)P(BlazersWon) \\
P(DieRoll_1, DieRoll_2) &= P(DieRoll_1)P(DieRoll_2)
\end{aligned}
$$

Independence

Key Ideas

- Sometimes no connection exists between variables
- Such independence determined by world knowledge

Formal Independence

$P(a, b) = P(a)P(b)$ or $P(a|b) = P(a)$

Examples

$$
\begin{aligned}
P(\mathit{EyeColor}, \mathit{Gender}) &= P(\mathit{EyeColor})P(\mathit{Gender}) \\
P(\mathit{Cavity}, \mathit{BlazersWon}) &= P(\mathit{Cavity})P(\mathit{BlazersWon}) \\
P(\mathit{DieRoll}_1, \mathit{DieRoll}_2) &= P(\mathit{DieRoll}_1)P(\mathit{DieRoll}_2)
\end{aligned}
$$

Independence

Using Independence

- No need to store entire joint probability table
- Can store several smaller independent tables

Independence

Using Independence

- No need to store entire joint probability table
- Can store several smaller independent tables

Examples

Table Size

Independence

Using Independence

- No need to store entire joint probability table
- Can store several smaller independent tables

Examples

Table Size

$\mathbf{P}(DieRoll_1, DieRoll_2)$

Independence

Using Independence

- No need to store entire joint probability table
- Can store several smaller independent tables

Examples

	Table Size
$\mathbf{P}(DieRoll_1, DieRoll_2)$	$6 \cdot 6 = 36$

Independence

Using Independence

- No need to store entire joint probability table
- Can store several smaller independent tables

Examples

	Table Size
$\mathbf{P}(DieRoll_1, DieRoll_2)$	$6 \cdot 6 = 36$
$\mathbf{P}(DieRoll_1)\mathbf{P}(DieRoll_2)$	

Independence

Using Independence

- No need to store entire joint probability table
- Can store several smaller independent tables

Examples

	Table Size
$\mathbf{P}(DieRoll_1, DieRoll_2)$	$6 \cdot 6 = 36$
$\mathbf{P}(DieRoll_1)\mathbf{P}(DieRoll_2)$	$6 + 6 = 12$

Independence

Using Independence

- No need to store entire joint probability table
- Can store several smaller independent tables

Examples

	Table Size
$\mathbf{P}(DieRoll_1, DieRoll_2)$	$6 \cdot 6 = 36$
$\mathbf{P}(DieRoll_1)\mathbf{P}(DieRoll_2)$	$6 + 6 = 12$
$\mathbf{P}(Age, Sex, BlazersWon)$	

Independence

Using Independence

- No need to store entire joint probability table
- Can store several smaller independent tables

Examples

	Table Size
$\mathbf{P}(DieRoll_1, DieRoll_2)$	$6 \cdot 6 = 36$
$\mathbf{P}(DieRoll_1)\mathbf{P}(DieRoll_2)$	$6 + 6 = 12$
$\mathbf{P}(Age, Sex, BlazersWon)$	$125 \cdot 2 \cdot 2 = 500$

Independence

Using Independence

- No need to store entire joint probability table
- Can store several smaller independent tables

Examples

	Table Size
$\mathbf{P}(DieRoll_1, DieRoll_2)$	$6 \cdot 6 = 36$
$\mathbf{P}(DieRoll_1)\mathbf{P}(DieRoll_2)$	$6 + 6 = 12$
$\mathbf{P}(Age, Sex, BlazersWon)$	$125 \cdot 2 \cdot 2 = 500$
$\mathbf{P}(Age, Sex)\mathbf{P}(BlazersWon)$	

Independence

Using Independence

- No need to store entire joint probability table
- Can store several smaller independent tables

Examples

	Table Size
$\mathbf{P}(DieRoll_1, DieRoll_2)$	$6 \cdot 6 = 36$
$\mathbf{P}(DieRoll_1)\mathbf{P}(DieRoll_2)$	$6 + 6 = 12$
$\mathbf{P}(Age, Sex, BlazersWon)$	$125 \cdot 2 \cdot 2 = 500$
$\mathbf{P}(Age, Sex)\mathbf{P}(BlazersWon)$	$125 \cdot 2 + 2 = 252$

Conditional Independence

Key Idea

Two variables can sometimes become independent after the value of a third variable is observed

Definition

A random variable X is conditionally independent of random variable Y given the random variable Z if:

$$\mathbf{P}(X, Y|Z) = \mathbf{P}(X|Z)\mathbf{P}(Y|Z)$$

or

$$\mathbf{P}(X|Y, Z) = \mathbf{P}(X|Z)$$

Conditional Independence Example

Is *GrayHair* indpendent of *Bifocals*?

No, we expect the two to often come together, e.g.:

$P(\textit{gray-hair}, \textit{bifocals}) > P(\textit{gray-hair})P(\textit{bifocals})$

Is *GreyHair* independent of *Bifocals* given *Age*?

Yes, the bifocals add nothing if we know the age, e.g.:

$P(\textit{gray-hair}|\textit{bifocals}, \textit{Age}=x) = P(\textit{gray-hair}|\textit{Age}=x)$

Conditional Independence Example

Is *GrayHair* indpendent of *Bifocals*?

No, we expect the two to often come together, e.g.:

$$P(\textit{gray-hair}, \textit{bifocals}) > P(\textit{gray-hair})P(\textit{bifocals})$$

Is *GreyHair* independent of *Bifocals* given *Age*?

Yes, the bifocals add nothing if we know the age, e.g.:

$$P(\textit{gray-hair} | \textit{bifocals}, \textit{Age}=x) = P(\textit{gray-hair} | \textit{Age}=x)$$

Conditional Independence Example

Is *GrayHair* indpendent of *Bifocals*?

No, we expect the two to often come together, e.g.:

$P(\textit{gray-hair}, \textit{bifocals}) > P(\textit{gray-hair})P(\textit{bifocals})$

Is *GreyHair* independent of *Bifocals* given *Age*?

Yes, the bifocals add nothing if we know the age, e.g.:

$P(\textit{gray-hair} | \textit{bifocals}, \textit{Age} = x) = P(\textit{gray-hair} | \textit{Age} = x)$

Conditional Independence Example

Is *GrayHair* indpendent of *Bifocals*?

No, we expect the two to often come together, e.g.:

$P(\textit{gray-hair}, \textit{bifocals}) > P(\textit{gray-hair})P(\textit{bifocals})$

Is *GreyHair* independent of *Bifocals* given *Age*?

Yes, the bifocals add nothing if we know the age, e.g.:

$P(\textit{gray-hair}|\textit{bifocals}, \textit{Age}=x) = P(\textit{gray-hair}|\textit{Age}=x)$

Conditional Independence Example

Noisy Phone

- Adam calls Betty and Charlie and says a number N_A
- Betty hears N_B and Charlie hears N_C

Are N_B and N_C independent?

No, we expect the numbers to be similar, e.g.:

$$P(N_B = 1, N_C = 1) > P(N_B = 1)P(N_C = 1)$$

Are N_B and N_C independent given N_A?

Yes, Betty's number adds nothing if we know Adam's, e.g.:

$$P(N_C = 1 | N_B = 2, N_A = 1) = P(N_C = 1 | N_A = 1)$$

Conditional Independence Example

Noisy Phone

- Adam calls Betty and Charlie and says a number N_A
- Betty hears N_B and Charlie hears N_C

Are N_B and N_C independent?

No, we expect the numbers to be similar, e.g.:

$$P(N_B = 1, N_C = 1) > P(N_B = 1)P(N_C = 1)$$

Are N_B and N_C independent given N_A?

Yes, Betty's number adds nothing if we know Adam's, e.g.:

$$P(N_C = 1 | N_B = 2, N_A = 1) = P(N_C = 1 | N_A = 1)$$

Conditional Independence Example

Noisy Phone

- Adam calls Betty and Charlie and says a number N_A
- Betty hears N_B and Charlie hears N_C

Are N_B and N_C independent?

No, we expect the numbers to be similar, e.g.:

$$P(N_B = 1, N_C = 1) > P(N_B = 1)P(N_C = 1)$$

Are N_B and N_C independent given N_A?

Yes, Betty's number adds nothing if we know Adam's, e.g.:

$$P(N_C = 1 | N_B = 2, N_A = 1) = P(N_C = 1 | N_A = 1)$$

Conditional Independence Example

Noisy Phone

- Adam calls Betty and Charlie and says a number N_A
- Betty hears N_B and Charlie hears N_C

Are N_B and N_C independent?

No, we expect the numbers to be similar, e.g.:

$$P(N_B = 1, N_C = 1) > P(N_B = 1)P(N_C = 1)$$

Are N_B and N_C independent given N_A?

Yes, Betty's number adds nothing if we know Adam's, e.g.:

$$P(N_C = 1 | N_B = 2, N_A = 1) = P(N_C = 1 | N_A = 1)$$

Conditional Independence Example

Noisy Phone

- Adam calls Betty and Charlie and says a number N_A
- Betty hears N_B and Charlie hears N_C

Are N_B and N_C independent?

No, we expect the numbers to be similar, e.g.:

$$P(N_B = 1, N_C = 1) > P(N_B = 1)P(N_C = 1)$$

Are N_B and N_C independent given N_A?

Yes, Betty's number adds nothing if we know Adam's, e.g.:

$$P(N_C = 1 | N_B = 2, N_A = 1) = P(N_C = 1 | N_A = 1)$$

Bayes' Rule

Key Idea

Swap the conditioned and conditioning variables

Bayes' Rule

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

Derivation

$$\begin{aligned} P(b|a) &= \frac{P(a \wedge b)}{P(a)} && \text{Definition} \\ &= \frac{P(a|b)P(b)}{P(a)} && \text{Product Rule} \end{aligned}$$

Bayes' Rule

Key Idea

Swap the conditioned and conditioning variables

Bayes' Rule

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

Derivation

$$\begin{aligned} P(b|a) &= \frac{P(a \wedge b)}{P(a)} && \text{Definition} \\ &= \frac{P(a|b)P(b)}{P(a)} && \text{Product Rule} \end{aligned}$$

Why Use Bayes' Rule?

Making Diagnoses Based on Causal Knowledge

$$P(Cause|Effect) = \frac{P(Effect|Cause)P(Cause)}{P(Effect)}$$

Example

$$P(meningitis|stiff\text{-}neck) = \frac{P(stiff\text{-}neck|meningitis)P(meningitis)}{P(stiff\text{-}neck)}$$

In an epidemic, where $P(meningitis)$ rises:

+Bayes $P(meningitis|stiff\text{-}neck)$ rises proportionally

-Bayes Collect data, re-estimate $P(meningitis|stiff\text{-}neck)$

Why Use Bayes' Rule?

Making Diagnoses Based on Causal Knowledge

$$P(\mathit{Cause}|\mathit{Effect}) = \frac{P(\mathit{Effect}|\mathit{Cause})P(\mathit{Cause})}{P(\mathit{Effect})}$$

Example

$$P(\mathit{meningitis}|\mathit{stiff\text{-}neck}) = \frac{P(\mathit{stiff\text{-}neck}|\mathit{meningitis})P(\mathit{meningitis})}{P(\mathit{stiff\text{-}neck})}$$

In an epidemic, where $P(\mathit{meningitis})$ rises:

+Bayes $P(\mathit{meningitis}|\mathit{stiff\text{-}neck})$ rises proportionally

-Bayes Collect data, re-estimate $P(\mathit{meningitis}|\mathit{stiff\text{-}neck})$

Why Use Bayes' Rule?

Making Diagnoses Based on Causal Knowledge

$$P(Cause|Effect) = \frac{P(Effect|Cause)P(Cause)}{P(Effect)}$$

Example

$$P(meningitis|stiff\text{-}neck) = \frac{P(stiff\text{-}neck|meningitis)P(meningitis)}{P(stiff\text{-}neck)}$$

In an epidemic, where $P(meningitis)$ rises:

+Bayes $P(meningitis|stiff\text{-}neck)$ rises proportionally

-Bayes Collect data, re-estimate $P(meningitis|stiff\text{-}neck)$

Bayes' Rule & Conditional Independence

Deriving More Manageable Models

$P(Cavity|toothache, catch)$

Bayes' Rule & Conditional Independence

Deriving More Manageable Models

$$\begin{aligned}
&\mathsf{P}(Cavity|toothache, catch) \\
&\quad = \alpha \mathsf{P}(toothache, catch|Cavity)\mathsf{P}(Cavity)
\end{aligned}$$

Bayes' Rule & Conditional Independence

Deriving More Manageable Models

$$\begin{aligned} &\mathsf{P}(Cavity|toothache, catch) \\ &\quad = \alpha \mathsf{P}(toothache, catch|Cavity)\mathsf{P}(Cavity) \\ &\quad = \alpha \mathsf{P}(toothache|Cavity)\mathsf{P}(catch|Cavity)\mathsf{P}(Cavity) \end{aligned}$$

Bayes' Rule & Conditional Independence

Deriving More Manageable Models

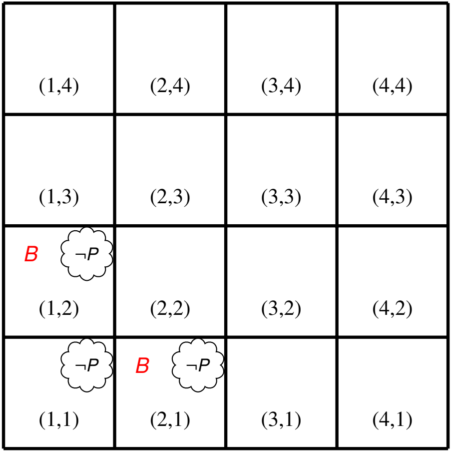
$$
\begin{aligned}
&P(Cavity|toothache, catch) \\
&\quad = \alpha P(toothache, catch|Cavity)P(Cavity) \\
&\quad = \alpha P(toothache|Cavity)P(catch|Cavity)P(Cavity)
\end{aligned}
$$

Naive Bayes Models

$P(Cause, Effect_1, \ldots, Effect_n) = P(Cause) \prod_i P(Effect_i|Cause)$

- All effects conditionally independent given cause
- Common class of machine learning models

Wumpus World Example



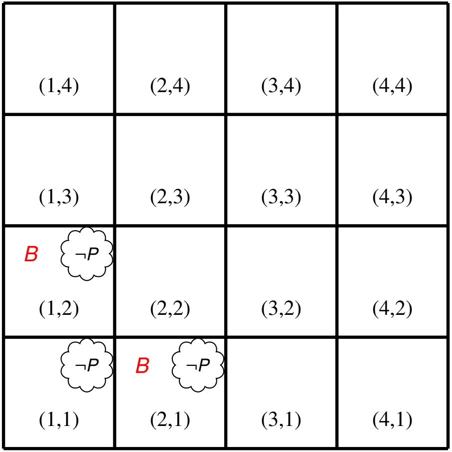
Query

$P_{1,3}$?

Brute Force Approach

- Calculate full joint distribution
- Sum all rows where: $p_{1,3}, \neg p_{1,1}, \neg p_{1,2}, \neg p_{2,1}, b_{1,2}, b_{2,1}$
- Result is $P(p_{1,3})$

Wumpus World Example



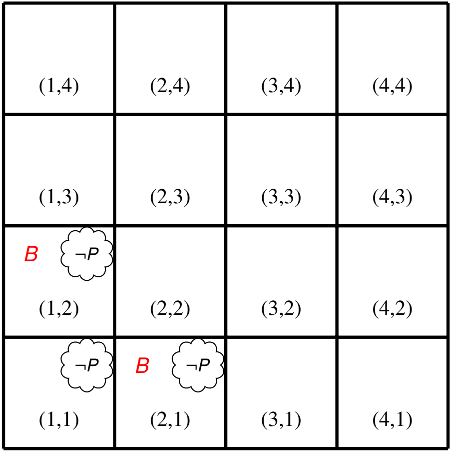
Query

$P_{1,3}$?

Brute Force Approach

- Calculate full joint distribution
- Sum all rows where: $p_{1,3}$, $\neg p_{1,1}$, $\neg p_{1,2}$, $\neg p_{2,1}$, $b_{1,2}$, $b_{2,1}$
- Result is $P(p_{1,3})$

Wumpus World Example



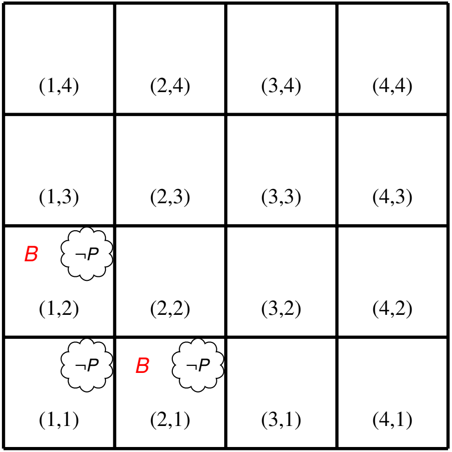
Query

$P_{1,3}$?

Brute Force Approach

- Calculate full joint distribution
- Sum all rows where: $p_{1,3}$, $\neg p_{1,1}$, $\neg p_{1,2}$, $\neg p_{2,1}$, $b_{1,2}$, $b_{2,1}$
- Result is $P(p_{1,3})$

Wumpus World Example



Query

$P_{1,3}$?

Brute Force Approach

- Calculate full joint distribution
- Sum all rows where: $p_{1,3}$, $\neg p_{1,1}$, $\neg p_{1,2}$, $\neg p_{2,1}$, $b_{1,2}$, $b_{2,1}$
- Result is $P(p_{1,3})$

Wumpus World Brute Force

(1,4)	(2,4)	(3,4)	(4,4)
(1,3)	(2,3)	(3,3)	(4,3)
(1,2)	(2,2)	(3,2)	(4,2)
(1,1)	(2,1)	(3,1)	(4,1)

Full Joint Distribution

$\mathbf{P}(P_{11}, \ldots, P_{44}, B_{11}, \ldots, B_{44})$

Total Rows?

$2^{16+16} = 2^{32}$

Rows Selected?

Assume:

$p_{13}, \neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}$

$2^{12+14} = 2^{28}$

There must be a better way!

Wumpus World Brute Force

(1,4)	(2,4)	(3,4)	(4,4)
(1,3)	(2,3)	(3,3)	(4,3)
(1,2)	(2,2)	(3,2)	(4,2)
(1,1)	(2,1)	(3,1)	(4,1)

Full Joint Distribution

$\mathbf{P}(P_{11}, \ldots, P_{44}, B_{11}, \ldots, B_{44})$

Total Rows?

$2^{16+16} = 2^{32}$

Rows Selected?

Assume:

$p_{13}, \neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}$

$2^{12+14} = 2^{28}$

There must be a better way!

Wumpus World Brute Force

(1,4)	(2,4)	(3,4)	(4,4)
(1,3)	(2,3)	(3,3)	(4,3)
(1,2)	(2,2)	(3,2)	(4,2)
(1,1)	(2,1)	(3,1)	(4,1)

Full Joint Distribution

$\mathbf{P}(P_{11}, \ldots, P_{44}, B_{11}, \ldots, B_{44})$

Total Rows?

$2^{16+16} = 2^{32}$

Rows Selected?

Assume:

$p_{13}, \neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}$

$2^{12+14} = 2^{28}$

There must be a better way!

Wumpus World Brute Force

(1,4)	(2,4)	(3,4)	(4,4)
(1,3)	(2,3)	(3,3)	(4,3)
(1,2)	(2,2)	(3,2)	(4,2)
(1,1)	(2,1)	(3,1)	(4,1)

Full Joint Distribution

$\mathbf{P}(P_{11}, \ldots, P_{44}, B_{11}, \ldots, B_{44})$

Total Rows?

$2^{16+16} = 2^{32}$

Rows Selected?

Assume:

$p_{13}, \neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}$

$2^{12+14} = 2^{28}$

There must be a better way!

Wumpus World Brute Force

(1,4)	(2,4)	(3,4)	(4,4)
(1,3)	(2,3)	(3,3)	(4,3)
(1,2)	(2,2)	(3,2)	(4,2)
(1,1)	(2,1)	(3,1)	(4,1)

Full Joint Distribution

$\mathbf{P}(P_{11}, \ldots, P_{44}, B_{11}, \ldots, B_{44})$

Total Rows?

$2^{16+16} = 2^{32}$

Rows Selected?

Assume:

$p_{13}, \neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}$

$2^{12+14} = 2^{28}$

There must be a better way!

Wumpus World Brute Force

(1,4)	(2,4)	(3,4)	(4,4)
(1,3)	(2,3)	(3,3)	(4,3)
(1,2)	(2,2)	(3,2)	(4,2)
(1,1)	(2,1)	(3,1)	(4,1)

Full Joint Distribution

$\mathbf{P}(P_{11}, \ldots, P_{44}, B_{11}, \ldots, B_{44})$

Total Rows?

$2^{16+16} = 2^{32}$

Rows Selected?

Assume:

$p_{13}, \neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}$

$2^{12+14} = 2^{28}$

There must be a better way!

Wumpus World Brute Force

(1,4)	(2,4)	(3,4)	(4,4)
(1,3)	(2,3)	(3,3)	(4,3)
(1,2)	(2,2)	(3,2)	(4,2)
(1,1)	(2,1)	(3,1)	(4,1)

Full Joint Distribution

$\mathbf{P}(P_{11}, \ldots, P_{44}, B_{11}, \ldots, B_{44})$

Total Rows?

$2^{16+16} = 2^{32}$

Rows Selected?

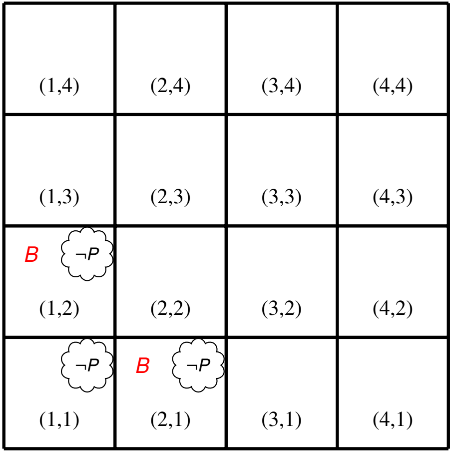
Assume:

$p_{13}, \neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}$

$2^{12+14} = 2^{28}$

There must be a better way!

Wumpus World with Independence



Key Idea

Observations conditionally independent of other squares given neighbors

Goal

Manipulate equation until we can apply conditional independence formula

$$\mathbf{P}(b_{12}, b_{21} | \neg p_{11}, \neg p_{12}, \neg p_{21}, P_{13}, P_{fringe}, P_{other}) =$$
$$\mathbf{P}(b_{12}, b_{21} | \neg p_{11}, \neg p_{12}, \neg p_{21}, P_{13}, P_{fringe})$$

Wumpus World with Independence

Other (1,4)	Other (2,4)	Other (3,4)	Other (4,4)
Query (1,3)	Other (2,3)	Other (3,3)	Other (4,3)
B $\neg P$ (1,2)	Fringe (2,2)	Other (3,2)	Other (4,2)
$\neg P$ (1,1)	B $\neg P$ (2,1)	Fringe (3,1)	Other (4,1)

Key Idea

Observations conditionally independent of other squares given neighbors

Goal

Manipulate equation until we can apply conditional independence formula

$$\mathbf{P}(b_{12}, b_{21} | \neg p_{11}, \neg p_{12}, \neg p_{21}, P_{13}, P_{fringe}, P_{other}) =$$
$$\mathbf{P}(b_{12}, b_{21} | \neg p_{11}, \neg p_{12}, \neg p_{21}, P_{13}, P_{fringe})$$

Wumpus World with Independence

Other (1,4)	Other (2,4)	Other (3,4)	Other (4,4)
Query (1,3)	Other (2,3)	Other (3,3)	Other (4,3)
B $\neg P$ (1,2)	Fringe (2,2)	Other (3,2)	Other (4,2)
$\neg P$ (1,1)	B $\neg P$ (2,1)	Fringe (3,1)	Other (4,1)

Key Idea

Observations conditionally independent of other squares given neighbors

Goal

Manipulate equation until we can apply conditional independence formula

$$\mathbf{P}(b_{12}, b_{21} | \neg p_{11}, \neg p_{12}, \neg p_{21}, P_{13}, P_{fringe}, P_{other}) =$$
$$\mathbf{P}(b_{12}, b_{21} | \neg p_{11}, \neg p_{12}, \neg p_{21}, P_{13}, P_{fringe})$$

Wumpus World with Independence

Other (1,4)	Other (2,4)	Other (3,4)	Other (4,4)
Query (1,3)	Other (2,3)	Other (3,3)	Other (4,3)
B $\neg P$ (1,2)	Fringe (2,2)	Other (3,2)	Other (4,2)
$\neg P$ (1,1)	B $\neg P$ (2,1)	Fringe (3,1)	Other (4,1)

Key Idea

Observations conditionally independent of other squares given neighbors

Goal

Manipulate equation until we can apply conditional independence formula

$$\mathbf{P}(b_{12}, b_{21} | \neg p_{11}, \neg p_{12}, \neg p_{21}, P_{13}, P_{fringe}, P_{other}) =$$
$$\mathbf{P}(b_{12}, b_{21} | \neg p_{11}, \neg p_{12}, \neg p_{21}, P_{13}, P_{fringe})$$

Wumpus World Equation Manipulation

$P(p_{13}|\neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21})$

Wumpus World Equation Manipulation

$$
\begin{aligned}
& P(p_{13} | \neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
& \quad = \quad \alpha P(p_{13}, \neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21})
\end{aligned}
$$

Wumpus World Equation Manipulation

$$
\begin{aligned}
& P(p_{13} | \neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
& \quad = \quad \alpha P(p_{13}, \neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
& \quad = \quad \alpha P(p_{13}, p_{known}, b_{12}, b_{21})
\end{aligned}
$$

Wumpus World Equation Manipulation

$$\begin{aligned}
&P(p_{13}|\neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
&\quad = \alpha P(p_{13}, \neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
&\quad = \alpha P(p_{13}, p_{known}, b_{12}, b_{21}) \\
&\quad = \alpha \sum_{fringe+other} P(p_{13}, p_{known}, p_{fringe+other}, b_{12}, b_{21})
\end{aligned}$$

Wumpus World Equation Manipulation

$$
\begin{aligned}
&P(p_{13} | \neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
&\quad = \alpha P(p_{13}, \neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
&\quad = \alpha P(p_{13}, p_{known}, b_{12}, b_{21}) \\
&\quad = \alpha \sum_{fringe+other} P(p_{13}, p_{known}, p_{fringe+other}, b_{12}, b_{21}) \\
&\quad = \alpha \sum_{fringe} \sum_{other} P(p_{13}, p_{known}, p_{fringe}, p_{other}, b_{12}, b_{21})
\end{aligned}
$$

Wumpus World Equation Manipulation

$$
\begin{aligned}
&P(p_{13} \mid \neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
&\quad = \alpha P(p_{13}, \neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
&\quad = \alpha P(p_{13}, p_{known}, b_{12}, b_{21}) \\
&\quad = \alpha \sum_{fringe+other} P(p_{13}, p_{known}, p_{fringe+other}, b_{12}, b_{21}) \\
&\quad = \alpha \sum_{fringe} \sum_{other} P(p_{13}, p_{known}, p_{fringe}, p_{other}, b_{12}, b_{21}) \\
&\quad = \alpha \sum_{fringe} \sum_{other} P(b_{12}, b_{21} \mid p_{13}, p_{known}, p_{fringe}, p_{other}) P(p_{13}, p_{known}, p_{fringe}, p_{other})
\end{aligned}
$$

Wumpus World Equation Manipulation

$$
\begin{aligned}
& P(p_{13} | \neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
& \quad = \alpha P(p_{13}, \neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
& \quad = \alpha P(p_{13}, p_{known}, b_{12}, b_{21}) \\
& \quad = \alpha \sum_{fringe+other} P(p_{13}, p_{known}, p_{fringe+other}, b_{12}, b_{21}) \\
& \quad = \alpha \sum_{fringe} \sum_{other} P(p_{13}, p_{known}, p_{fringe}, p_{other}, b_{12}, b_{21}) \\
& \quad = \alpha \sum_{fringe} \sum_{other} P(b_{12}, b_{21} | p_{13}, p_{known}, p_{fringe}, p_{other}) P(p_{13}, p_{known}, p_{fringe}, p_{other}) \\
& \quad = \alpha \sum_{fringe} \sum_{other} P(b_{12}, b_{21} | p_{13}, p_{known}, p_{fringe}) P(p_{13}, p_{known}, p_{fringe}, p_{other})
\end{aligned}
$$

Wumpus World Equation Manipulation

$$
\begin{aligned}
& P(p_{13} | \neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
& \quad = \alpha P(p_{13}, \neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
& \quad = \alpha P(p_{13}, p_{known}, b_{12}, b_{21}) \\
& \quad = \alpha \sum_{fringe+other} P(p_{13}, p_{known}, p_{fringe+other}, b_{12}, b_{21}) \\
& \quad = \alpha \sum_{fringe} \sum_{other} P(p_{13}, p_{known}, p_{fringe}, p_{other}, b_{12}, b_{21}) \\
& \quad = \alpha \sum_{fringe} \sum_{other} P(b_{12}, b_{21} | p_{13}, p_{known}, p_{fringe}, p_{other}) P(p_{13}, p_{known}, p_{fringe}, p_{other}) \\
& \quad = \alpha \sum_{fringe} \sum_{other} P(b_{12}, b_{21} | p_{13}, p_{known}, p_{fringe}) P(p_{13}, p_{known}, p_{fringe}, p_{other}) \\
& \quad = \alpha \sum_{fringe} \sum_{other} P(b_{12}, b_{21} | p_{13}, p_{known}, p_{fringe}) P(p_{13}) P(p_{known}) P(p_{fringe}) P(p_{other})
\end{aligned}
$$

Wumpus World Equation Manipulation

$$
\begin{aligned}
& P(p_{13}|\neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
& \quad = \quad \alpha P(p_{13}, \neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
& \quad = \quad \alpha P(p_{13}, p_{known}, b_{12}, b_{21}) \\
& \quad = \quad \alpha \sum_{fringe+other} P(p_{13}, p_{known}, p_{fringe+other}, b_{12}, b_{21}) \\
& \quad = \quad \alpha \sum_{fringe} \sum_{other} P(p_{13}, p_{known}, p_{fringe}, p_{other}, b_{12}, b_{21}) \\
& \quad = \quad \alpha \sum_{fringe} \sum_{other} P(b_{12}, b_{21}|p_{13}, p_{known}, p_{fringe}, p_{other}) P(p_{13}, p_{known}, p_{fringe}, p_{other}) \\
& \quad = \quad \alpha \sum_{fringe} \sum_{other} P(b_{12}, b_{21}|p_{13}, p_{known}, p_{fringe}) P(p_{13}, p_{known}, p_{fringe}, p_{other}) \\
& \quad = \quad \alpha \sum_{fringe} \sum_{other} P(b_{12}, b_{21}|p_{13}, p_{known}, p_{fringe}) P(p_{13}) P(p_{known}) P(p_{fringe}) P(p_{other}) \\
& \quad = \quad \alpha \sum_{fringe} P(b_{12}, b_{21}|p_{13}, p_{known}, p_{fringe}) P(p_{13}) P(p_{known}) P(p_{fringe}) \sum_{other} P(p_{other})
\end{aligned}
$$

Wumpus World Equation Manipulation

$$
\begin{aligned}
& P(p_{13}|\neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
& = \alpha P(p_{13}, \neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
& = \alpha P(p_{13}, p_{known}, b_{12}, b_{21}) \\
& = \alpha \sum_{fringe+other} P(p_{13}, p_{known}, p_{fringe+other}, b_{12}, b_{21}) \\
& = \alpha \sum_{fringe} \sum_{other} P(p_{13}, p_{known}, p_{fringe}, p_{other}, b_{12}, b_{21}) \\
& = \alpha \sum_{fringe} \sum_{other} P(b_{12}, b_{21}|p_{13}, p_{known}, p_{fringe}, p_{other}) P(p_{13}, p_{known}, p_{fringe}, p_{other}) \\
& = \alpha \sum_{fringe} \sum_{other} P(b_{12}, b_{21}|p_{13}, p_{known}, p_{fringe}) P(p_{13}, p_{known}, p_{fringe}, p_{other}) \\
& = \alpha \sum_{fringe} \sum_{other} P(b_{12}, b_{21}|p_{13}, p_{known}, p_{fringe}) P(p_{13}) P(p_{known}) P(p_{fringe}) P(p_{other}) \\
& = \alpha \sum_{fringe} P(b_{12}, b_{21}|p_{13}, p_{known}, p_{fringe}) P(p_{13}) P(p_{known}) P(p_{fringe}) \sum_{other} P(p_{other}) \\
& = \alpha \sum_{fringe} P(b_{12}, b_{21}|p_{13}, p_{known}, p_{fringe}) P(p_{13}) P(p_{known}) P(p_{fringe})
\end{aligned}
$$

Wumpus World Equation Manipulation

$$
\begin{aligned}
& P(p_{13} | \neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
& \quad = \alpha P(p_{13}, \neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
& \quad = \alpha P(p_{13}, p_{known}, b_{12}, b_{21}) \\
& \quad = \alpha \sum_{fringe+other} P(p_{13}, p_{known}, p_{fringe+other}, b_{12}, b_{21}) \\
& \quad = \alpha \sum_{fringe} \sum_{other} P(p_{13}, p_{known}, p_{fringe}, p_{other}, b_{12}, b_{21}) \\
& \quad = \alpha \sum_{fringe} \sum_{other} P(b_{12}, b_{21} | p_{13}, p_{known}, p_{fringe}, p_{other}) P(p_{13}, p_{known}, p_{fringe}, p_{other}) \\
& \quad = \alpha \sum_{fringe} \sum_{other} P(b_{12}, b_{21} | p_{13}, p_{known}, p_{fringe}) P(p_{13}, p_{known}, p_{fringe}, p_{other}) \\
& \quad = \alpha \sum_{fringe} \sum_{other} P(b_{12}, b_{21} | p_{13}, p_{known}, p_{fringe}) P(p_{13}) P(p_{known}) P(p_{fringe}) P(p_{other}) \\
& \quad = \alpha \sum_{fringe} P(b_{12}, b_{21} | p_{13}, p_{known}, p_{fringe}) P(p_{13}) P(p_{known}) P(p_{fringe}) \sum_{other} P(p_{other}) \\
& \quad = \alpha \sum_{fringe} P(b_{12}, b_{21} | p_{13}, p_{known}, p_{fringe}) P(p_{13}) P(p_{known}) P(p_{fringe}) \\
& \quad = \alpha P(p_{13}) P(p_{known}) \sum_{fringe} P(b_{12}, b_{21} | p_{13}, p_{known}, p_{fringe}) P(p_{fringe})
\end{aligned}
$$

Wumpus World Probability Calculation

Goal: Calculate the Summation

$$\sum_{p_{22},p_{31}} P(b_{12}, b_{21} | p_{13}, \neg p_{11}, \neg p_{12}, \neg p_{21}, P_{22}, P_{31}) P(P_{22}, P_{31})$$

Key Ideas

- First term is 1 if observed breezes are beside pits
- First term is 0 if observed breezes are not beside pits
- So find worlds consistent with observations
- And sum $P(P_{22}, P_{31})$ over these worlds

Wumpus World Probability Calculation

Goal: Calculate the Summation

$$\sum_{p_{22},p_{31}} P(b_{12}, b_{21} | p_{13}, \neg p_{11}, \neg p_{12}, \neg p_{21}, P_{22}, P_{31}) P(P_{22}, P_{31})$$

Key Ideas

- First term is 1 if observed breezes are beside pits
- First term is 0 if observed breezes are not beside pits
- So find worlds consistent with observations
- And sum $P(P_{22}, P_{31})$ over these worlds

Wumpus World Probability Calculation

Goal: Calculate the Summation

$$\sum_{p_{22},p_{31}} P(b_{12}, b_{21} | p_{13}, \neg p_{11}, \neg p_{12}, \neg p_{21}, P_{22}, P_{31}) P(P_{22}, P_{31})$$

Key Ideas

- First term is 1 if observed breezes are beside pits
- First term is 0 if observed breezes are not beside pits
- So find worlds consistent with observations
- And sum $P(P_{22}, P_{31})$ over these worlds

Wumpus World Probability Calculation

Worlds Consistent with Observations

P		
B *¬P*	**P**	
¬P	*B* *¬P*	**P**

P		
B *¬P*	**P**	
¬P	*B* *¬P*	**¬P**

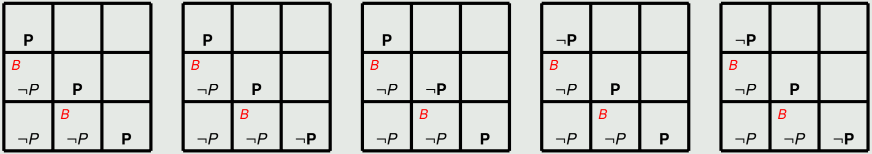
P		
B *¬P*	**¬P**	
¬P	*B* *¬P*	**P**

¬P		
B *¬P*	**P**	
¬P	*B* *¬P*	**P**

¬P		
B *¬P*	**P**	
¬P	*B* *¬P*	**¬P**

Wumpus World Probability Calculation

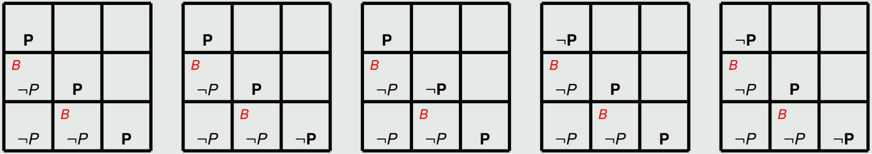
Worlds Consistent with Observations



Calculating $P(P_{22}, P_{31})$, assuming pit probability is 0.2

Wumpus World Probability Calculation

Worlds Consistent with Observations



$.2 \times .2 = .04$

Calculating $P(P_{22}, P_{31})$, assuming pit probability is 0.2

Wumpus World Probability Calculation

Worlds Consistent with Observations

P		
B ¬*P*	**P**	
¬*P*	*B* ¬*P*	**P**

$.2 \times .2 = .04$

P		
B ¬*P*	**P**	
¬*P*	*B* ¬*P*	¬**P**

$.2 \times .8 = .16$

P		
B ¬*P*	¬**P**	
¬*P*	*B* ¬*P*	**P**

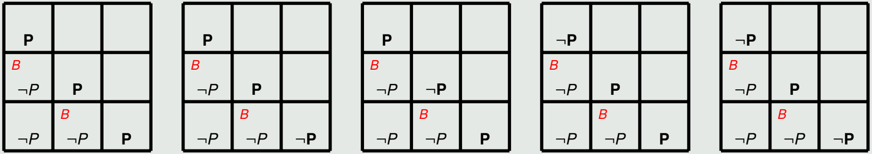
¬**P**		
B ¬*P*	**P**	
¬*P*	*B* ¬*P*	**P**

¬**P**		
B ¬*P*	**P**	
¬*P*	*B* ¬*P*	¬**P**

Calculating $P(P_{22}, P_{31})$, assuming pit probability is 0.2

Wumpus World Probability Calculation

Worlds Consistent with Observations



$.2 \times .2 = .04$ $.2 \times .8 = .16$ $.8 \times .2 = .16$

Calculating $P(P_{22}, P_{31})$, assuming pit probability is 0.2

Wumpus World Probability Calculation

Worlds Consistent with Observations

P		
B $\neg P$	**P**	
$\neg P$	*B* $\neg P$	**P**

P		
B $\neg P$	**P**	
$\neg P$	*B* $\neg P$	**¬P**

P		
B $\neg P$	**¬P**	
$\neg P$	*B* $\neg P$	**P**

¬P		
B $\neg P$	**P**	
$\neg P$	*B* $\neg P$	**P**

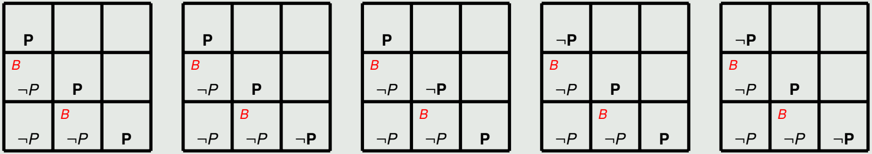
¬P		
B $\neg P$	**P**	
$\neg P$	*B* $\neg P$	**¬P**

$.2 \times .2 = .04$ $.2 \times .8 = .16$ $.8 \times .2 = .16$ $.2 \times .2 = .04$

Calculating $P(P_{22}, P_{31})$, assuming pit probability is 0.2

Wumpus World Probability Calculation

Worlds Consistent with Observations

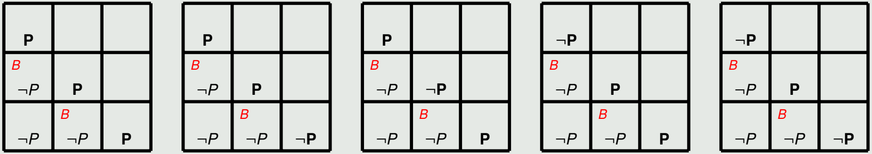


$.2 \times .2 = .04$ $.2 \times .8 = .16$ $.8 \times .2 = .16$ $.2 \times .2 = .04$ $.2 \times .8 = .16$

Calculating $P(P_{22}, P_{31})$, assuming pit probability is 0.2

Wumpus World Probability Calculation

Worlds Consistent with Observations



$.2 \times .2 = .04$ $.2 \times .8 = .16$ $.8 \times .2 = .16$ $.2 \times .2 = .04$ $.2 \times .8 = .16$

Calculating $P(P_{22}, P_{31})$, assuming pit probability is 0.2

Final Calculation

$$\sum_{p_{22},p_{31}} P(b_{12}, b_{21} | p_{13}, \ldots) P(P_{22}, P_{31}) = 0.04 + 0.16 + 0.16$$

$$\sum_{p_{22},p_{31}} P(b_{12}, b_{21} | \neg p_{13}, \ldots) P(P_{22}, P_{31}) = 0.04 + 0.16$$

Wumpus World Probability Calculation

$P(p_{13}|\neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21})$

Wumpus World Probability Calculation

$$
\begin{aligned}
& P(p_{13}|\neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
& \quad = \ \alpha P(p_{13}) P(\neg p_{11}, \neg p_{12}, \neg p_{21}) \sum \ldots p_{13} \ldots
\end{aligned}
$$

Wumpus World Probability Calculation

$$
\begin{aligned}
& P(p_{13}|\neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
& \quad = \alpha P(p_{13}) P(\neg p_{11}, \neg p_{12}, \neg p_{21}) \sum \ldots p_{13} \ldots \\
& \quad = \alpha \times 0.2 \times P(\neg p_{11}, \neg p_{12}, \neg p_{21}) \sum \ldots p_{13} \ldots
\end{aligned}
$$

Wumpus World Probability Calculation

$$
\begin{aligned}
&P(p_{13}|\neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
&\quad = \alpha P(p_{13}) P(\neg p_{11}, \neg p_{12}, \neg p_{21}) \sum \ldots p_{13} \ldots \\
&\quad = \alpha \times 0.2 \times P(\neg p_{11}, \neg p_{12}, \neg p_{21}) \sum \ldots p_{13} \ldots \\
&\quad = \alpha \times 0.2 \times 0.8 \times 0.8 \times 0.8 \times \sum \ldots p_{13} \ldots
\end{aligned}
$$

Wumpus World Probability Calculation

$$
\begin{aligned}
& P(p_{13} | \neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
& \quad = \alpha P(p_{13}) P(\neg p_{11}, \neg p_{12}, \neg p_{21}) \sum \ldots p_{13} \ldots \\
& \quad = \alpha \times 0.2 \times P(\neg p_{11}, \neg p_{12}, \neg p_{21}) \sum \ldots p_{13} \ldots \\
& \quad = \alpha \times 0.2 \times 0.8 \times 0.8 \times 0.8 \times \sum \ldots p_{13} \ldots \\
& \quad = \alpha \times 0.2 \times 0.8 \times 0.8 \times 0.8 \times (0.04 + 0.16 + 0.16)
\end{aligned}
$$

Wumpus World Probability Calculation

$$
\begin{aligned}
& P(p_{13}|\neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
&= \alpha P(p_{13}) P(\neg p_{11}, \neg p_{12}, \neg p_{21}) \sum \ldots p_{13} \ldots \\
&= \alpha \times 0.2 \times P(\neg p_{11}, \neg p_{12}, \neg p_{21}) \sum \ldots p_{13} \ldots \\
&= \alpha \times 0.2 \times 0.8 \times 0.8 \times 0.8 \times \sum \ldots p_{13} \ldots \\
&= \alpha \times 0.2 \times 0.8 \times 0.8 \times 0.8 \times (0.04 + 0.16 + 0.16) \\
&= 0.036864\alpha
\end{aligned}
$$

Wumpus World Probability Calculation

$$
\begin{aligned}
& P(p_{13}|\neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
& \quad = \alpha P(p_{13}) P(\neg p_{11}, \neg p_{12}, \neg p_{21}) \sum \ldots p_{13} \ldots \\
& \quad = \alpha \times 0.2 \times P(\neg p_{11}, \neg p_{12}, \neg p_{21}) \sum \ldots p_{13} \ldots \\
& \quad = \alpha \times 0.2 \times 0.8 \times 0.8 \times 0.8 \times \sum \ldots p_{13} \ldots \\
& \quad = \alpha \times 0.2 \times 0.8 \times 0.8 \times 0.8 \times (0.04 + 0.16 + 0.16) \\
& \quad = 0.036864\alpha
\end{aligned}
$$

$$P(\neg p_{13}|\neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21})$$

Wumpus World Probability Calculation

$$
\begin{aligned}
&P(p_{13}|\neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
&\quad = \alpha P(p_{13}) P(\neg p_{11}, \neg p_{12}, \neg p_{21}) \textstyle\sum \ldots p_{13} \ldots \\
&\quad = \alpha \times 0.2 \times P(\neg p_{11}, \neg p_{12}, \neg p_{21}) \textstyle\sum \ldots p_{13} \ldots \\
&\quad = \alpha \times 0.2 \times 0.8 \times 0.8 \times 0.8 \times \textstyle\sum \ldots p_{13} \ldots \\
&\quad = \alpha \times 0.2 \times 0.8 \times 0.8 \times 0.8 \times (0.04 + 0.16 + 0.16) \\
&\quad = 0.036864\alpha
\end{aligned}
$$

$$
\begin{aligned}
&P(\neg p_{13}|\neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
&\quad = \alpha P(\neg p_{13}) P(\neg p_{11}, \neg p_{12}, \neg p_{21}) \textstyle\sum \ldots \neg p_{13} \ldots
\end{aligned}
$$

Wumpus World Probability Calculation

$$
\begin{aligned}
&P(p_{13}|\neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
&\quad = \alpha P(p_{13}) P(\neg p_{11}, \neg p_{12}, \neg p_{21}) \sum \ldots p_{13} \ldots \\
&\quad = \alpha \times 0.2 \times P(\neg p_{11}, \neg p_{12}, \neg p_{21}) \sum \ldots p_{13} \ldots \\
&\quad = \alpha \times 0.2 \times 0.8 \times 0.8 \times 0.8 \times \sum \ldots p_{13} \ldots \\
&\quad = \alpha \times 0.2 \times 0.8 \times 0.8 \times 0.8 \times (0.04 + 0.16 + 0.16) \\
&\quad = 0.036864\alpha
\end{aligned}
$$

$$
\begin{aligned}
&P(\neg p_{13}|\neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
&\quad = \alpha P(\neg p_{13}) P(\neg p_{11}, \neg p_{12}, \neg p_{21}) \sum \ldots \neg p_{13} \ldots \\
&\quad = \alpha \times 0.8 \times 0.8 \times 0.8 \times 0.8 \times (0.04 + 0.16)
\end{aligned}
$$

Wumpus World Probability Calculation

$$
\begin{aligned}
&P(p_{13}|\neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
&\quad = \alpha P(p_{13}) P(\neg p_{11}, \neg p_{12}, \neg p_{21}) \textstyle\sum \ldots p_{13} \ldots \\
&\quad = \alpha \times 0.2 \times P(\neg p_{11}, \neg p_{12}, \neg p_{21}) \textstyle\sum \ldots p_{13} \ldots \\
&\quad = \alpha \times 0.2 \times 0.8 \times 0.8 \times 0.8 \times \textstyle\sum \ldots p_{13} \ldots \\
&\quad = \alpha \times 0.2 \times 0.8 \times 0.8 \times 0.8 \times (0.04 + 0.16 + 0.16) \\
&\quad = 0.036864\alpha
\end{aligned}
$$

$$
\begin{aligned}
&P(\neg p_{13}|\neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
&\quad = \alpha P(\neg p_{13}) P(\neg p_{11}, \neg p_{12}, \neg p_{21}) \textstyle\sum \ldots \neg p_{13} \ldots \\
&\quad = \alpha \times 0.8 \times 0.8 \times 0.8 \times 0.8 \times (0.04 + 0.16) \\
&\quad = 0.08192\alpha
\end{aligned}
$$

Wumpus World Probability Calculation

$$
\begin{aligned}
&P(p_{13}|\neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
&\quad = \alpha P(p_{13}) P(\neg p_{11}, \neg p_{12}, \neg p_{21}) \sum \ldots p_{13} \ldots \\
&\quad = \alpha \times 0.2 \times P(\neg p_{11}, \neg p_{12}, \neg p_{21}) \sum \ldots p_{13} \ldots \\
&\quad = \alpha \times 0.2 \times 0.8 \times 0.8 \times 0.8 \times \sum \ldots p_{13} \ldots \\
&\quad = \alpha \times 0.2 \times 0.8 \times 0.8 \times 0.8 \times (0.04 + 0.16 + 0.16) \\
&\quad = 0.036864\alpha
\end{aligned}
$$

$$
\begin{aligned}
&P(\neg p_{13}|\neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
&\quad = \alpha P(\neg p_{13}) P(\neg p_{11}, \neg p_{12}, \neg p_{21}) \sum \ldots \neg p_{13} \ldots \\
&\quad = \alpha \times 0.8 \times 0.8 \times 0.8 \times 0.8 \times (0.04 + 0.16) \\
&\quad = 0.08192\alpha
\end{aligned}
$$

$$P(p_{13}|\neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21})$$

Wumpus World Probability Calculation

$$
\begin{aligned}
&P(p_{13}|\neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
&\quad = \alpha P(p_{13}) P(\neg p_{11}, \neg p_{12}, \neg p_{21}) \sum \ldots p_{13} \ldots \\
&\quad = \alpha \times 0.2 \times P(\neg p_{11}, \neg p_{12}, \neg p_{21}) \sum \ldots p_{13} \ldots \\
&\quad = \alpha \times 0.2 \times 0.8 \times 0.8 \times 0.8 \times \sum \ldots p_{13} \ldots \\
&\quad = \alpha \times 0.2 \times 0.8 \times 0.8 \times 0.8 \times (0.04 + 0.16 + 0.16) \\
&\quad = 0.036864\alpha
\end{aligned}
$$

$$
\begin{aligned}
&P(\neg p_{13}|\neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
&\quad = \alpha P(\neg p_{13}) P(\neg p_{11}, \neg p_{12}, \neg p_{21}) \sum \ldots \neg p_{13} \ldots \\
&\quad = \alpha \times 0.8 \times 0.8 \times 0.8 \times 0.8 \times (0.04 + 0.16) \\
&\quad = 0.08192\alpha
\end{aligned}
$$

$$
\begin{aligned}
&P(p_{13}|\neg p_{11}, \neg p_{12}, \neg p_{21}, b_{12}, b_{21}) \\
&\quad = 0.036864/(0.036864 + 0.08192) \approx 0.31
\end{aligned}
$$

Probability Rules

Product Rule

$$P(a, b) = P(a|b)P(b)$$

Independence

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

Conditional Independence

$$P(a, b|c) = P(a|c)P(b|c)$$
$$P(a|b, c) = P(a|c)$$

Probability Rules

Product Rule

$$P(a, b) = P(a|b)P(b)$$

Independence

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

Conditional Independence

$$P(a, b|c) = P(a|c)P(b|c)$$
$$P(a|b, c) = P(a|c)$$

Given: w indep. x, y, z
x indep. z given y

Show:

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

Probability Rules

Product Rule

$P(a, b) = P(a|b)P(b)$

Independence

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

Conditional Independence

$P(a, b|c) = P(a|c)P(b|c)$

$P(a|b, c) = P(a|c)$

Given: w indep. x, y, z

x indep. z given y

Show:

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

Proof:

$P(w, x, y, z)$

Probability Rules

Product Rule

$P(a, b) = P(a|b)P(b)$

Independence

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

Conditional Independence

$P(a, b|c) = P(a|c)P(b|c)$
$P(a|b, c) = P(a|c)$

Given: w indep. x, y, z
x indep. z given y

Show:

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

Proof:

$P(w, x, y, z)$
$= P(w)P(x, y, z)$

Probability Rules

Product Rule

$$P(a, b) = P(a|b)P(b)$$

Independence

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

Conditional Independence

$$P(a, b|c) = P(a|c)P(b|c)$$
$$P(a|b, c) = P(a|c)$$

Given: w indep. x, y, z
x indep. z given y

Show:

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

Proof:

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

Probability Rules

Product Rule

$P(a,b) = P(a|b)P(b)$

Independence

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

Conditional Independence

$P(a,b|c) = P(a|c)P(b|c)$
$P(a|b,c) = P(a|c)$

Given: w indep. x, y, z
x indep. z given y

Show:

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

Proof:

$$
\begin{aligned}
&P(w,x,y,z) \\
&= P(w)P(x,y,z) \\
&= P(w)P(x,z|y)P(y) \\
&= P(w)P(x|y)P(z|y)P(y)
\end{aligned}
$$

Probability Rules

Product Rule

$P(a, b) = P(a|b)P(b)$

Independence

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

Conditional Independence

$P(a, b|c) = P(a|c)P(b|c)$

$P(a|b, c) = P(a|c)$

Given: w indep. x, y, z

x indep. z given y

Show:

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

Proof:

$$\begin{aligned} & P(w, x, y, z) \\ &= P(w)P(x, y, z) \\ &= P(w)P(x, z|y)P(y) \\ &= P(w)P(x|y)P(z|y)P(y) \\ &= P(w)P(x|y)P(z, y) \end{aligned}$$

Probability Rule Exercises

Product Rule

$P(a, b) = P(a|b)P(b)$

Independence

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

Conditional Independence

$P(a, b|c) = P(a|c)P(b|c)$
$P(a|b, c) = P(a|c)$

Given:
z indep. w, x, y
w indep. x given y

Show:

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

Given:
w indep. y, z given x
x indep. z given y
y indep. z

Show:

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

Normalization Rules

Complements

$P(\neg a) + P(a) = 1$

Summing Out

$P(a, b) + P(\neg a, b) = P(b)$

Distributions

$$\sum_{i}^{d} P(X = x_i) = 1$$

Normalization Rules

Complements

$$P(\neg a) + P(a) = 1$$

Summing Out

$$P(a, b) + P(\neg a, b) = P(b)$$

Distributions

$$\sum_{i}^{d} P(X = x_i) = 1$$

Given: $\mathbf{P}(X|y, z) = \alpha\langle 0.1, 0.3\rangle$

Show: $\mathbf{P}(X|y, z) = \langle 0.25, 0.75\rangle$

Normalization Rules

Given: $\mathbf{P}(X|y,z) = \alpha\langle 0.1, 0.3\rangle$

Show: $\mathbf{P}(X|y,z) = \langle 0.25, 0.75\rangle$

$$P(y,z) = P(x,y,z) + P(\neg x,y,z)$$

Complements

$$P(\neg a) + P(a) = 1$$

Summing Out

$$P(a,b) + P(\neg a,b) = P(b)$$

Distributions

$$\sum_i^d P(X = x_i) = 1$$

Normalization Rules

Complements

$P(\neg a) + P(a) = 1$

Summing Out

$P(a, b) + P(\neg a, b) = P(b)$

Distributions

$$\sum_i^d P(X = x_i) = 1$$

Given: $\mathbf{P}(X|y, z) = \alpha\langle 0.1, 0.3\rangle$

Show: $\mathbf{P}(X|y, z) = \langle 0.25, 0.75\rangle$

$$P(y, z) = P(x, y, z) + P(\neg x, y, z)$$

$$1 = \frac{P(x, y, z) + P(\neg x, y, c)}{P(y, z)}$$

Normalization Rules

Complements

$P(\neg a) + P(a) = 1$

Summing Out

$P(a, b) + P(\neg a, b) = P(b)$

Distributions

$$\sum_{i}^{d} P(X = x_i) = 1$$

Given: $\mathbf{P}(X|y, z) = \alpha\langle 0.1, 0.3\rangle$

Show: $\mathbf{P}(X|y, z) = \langle 0.25, 0.75\rangle$

$$P(y, z) = P(x, y, z) + P(\neg x, y, z)$$

$$1 = \frac{P(x, y, z) + P(\neg x, y, c)}{P(y, z)}$$

$$1 = \frac{P(x, y, z)}{P(y, z)} + \frac{P(\neg x, y, z)}{P(y, z)}$$

Normalization Rules

Complements

$$P(\neg a) + P(a) = 1$$

Summing Out

$$P(a, b) + P(\neg a, b) = P(b)$$

Distributions

$$\sum_i^d P(X = x_i) = 1$$

Given: $\mathbf{P}(X|y, z) = \alpha\langle 0.1, 0.3\rangle$

Show: $\mathbf{P}(X|y, z) = \langle 0.25, 0.75\rangle$

$$P(y, z) = P(x, y, z) + P(\neg x, y, z)$$

$$1 = \frac{P(x, y, z) + P(\neg x, y, c)}{P(y, z)}$$

$$1 = \frac{P(x, y, z)}{P(y, z)} + \frac{P(\neg x, y, z)}{P(y, z)}$$

$$1 = P(x|y, z) + P(\neg x|y, z)$$

Normalization Rules

Complements

$$P(\neg a) + P(a) = 1$$

Summing Out

$$P(a, b) + P(\neg a, b) = P(b)$$

Distributions

$$\sum_i^d P(X = x_i) = 1$$

Given: $\mathbf{P}(X|y, z) = \alpha\langle 0.1, 0.3\rangle$

Show: $\mathbf{P}(X|y, z) = \langle 0.25, 0.75\rangle$

$$P(y, z) = P(x, y, z) + P(\neg x, y, z)$$

$$1 = \frac{P(x, y, z) + P(\neg x, y, c)}{P(y, z)}$$

$$1 = \frac{P(x, y, z)}{P(y, z)} + \frac{P(\neg x, y, z)}{P(y, z)}$$

$$1 = P(x|y, z) + P(\neg x|y, z)$$

$$1 = 0.1\alpha + 0.3\alpha$$

Normalization Rules

Complements

$$P(\neg a) + P(a) = 1$$

Summing Out

$$P(a, b) + P(\neg a, b) = P(b)$$

Distributions

$$\sum_i^d P(X = x_i) = 1$$

Given: $\mathbf{P}(X|y,z) = \alpha\langle 0.1, 0.3\rangle$

Show: $\mathbf{P}(X|y,z) = \langle 0.25, 0.75\rangle$

$$P(y,z) = P(x,y,z) + P(\neg x,y,z)$$

$$1 = \frac{P(x,y,z) + P(\neg x,y,c)}{P(y,z)}$$

$$1 = \frac{P(x,y,z)}{P(y,z)} + \frac{P(\neg x,y,z)}{P(y,z)}$$

$$1 = P(x|y,z) + P(\neg x|y,z)$$

$$1 = 0.1\alpha + 0.3\alpha$$

$$1 = 0.4\alpha$$

Normalization Rules

Complements

$P(\neg a) + P(a) = 1$

Summing Out

$P(a, b) + P(\neg a, b) = P(b)$

Distributions

$$\sum_i^d P(X = x_i) = 1$$

Given: $\mathbf{P}(X|y, z) = \alpha\langle 0.1, 0.3\rangle$

Show: $\mathbf{P}(X|y, z) = \langle 0.25, 0.75\rangle$

$$\begin{aligned}
P(y, z) &= P(x, y, z) + P(\neg x, y, z) \\
1 &= \frac{P(x, y, z) + P(\neg x, y, c)}{P(y, z)} \\
1 &= \frac{P(x, y, z)}{P(y, z)} + \frac{P(\neg x, y, z)}{P(y, z)} \\
1 &= P(x|y, z) + P(\neg x|y, z) \\
1 &= 0.1\alpha + 0.3\alpha \\
1 &= 0.4\alpha \\
2.5 &= \alpha
\end{aligned}$$

Normalization Rules

Complements

$P(\neg a) + P(a) = 1$

Summing Out

$P(a, b) + P(\neg a, b) = P(b)$

Distributions

$$\sum_{i}^{d} P(X = x_i) = 1$$

Given: $\mathbf{P}(X|y, z) = \alpha\langle 0.1, 0.3\rangle$

Show: $\mathbf{P}(X|y, z) = \langle 0.25, 0.75\rangle$

$$P(y, z) = P(x, y, z) + P(\neg x, y, z)$$

$$1 = \frac{P(x, y, z) + P(\neg x, y, c)}{P(y, z)}$$

$$1 = \frac{P(x, y, z)}{P(y, z)} + \frac{P(\neg x, y, z)}{P(y, z)}$$

$$1 = P(x|y, z) + P(\neg x|y, z)$$

$$1 = 0.1\alpha + 0.3\alpha$$

$$1 = 0.4\alpha$$

$$2.5 = \alpha$$

Normalization Rule Practice

Complements

$$P(\neg a) + P(a) = 1$$
$$P(a|b) + P(\neg a|b) = 1$$

Summing Out

$$P(a, b) + P(\neg a, b) = P(b)$$

Distributions

$$\sum_i^d P(X = x_i) = 1$$

Given:

$$\begin{aligned} P(y|x) &= 0.4 \\ P(y|\neg x) &= 0.9 \\ P(x) &= 0.2 \\ P(\neg x) &= 0.8 \end{aligned}$$

Show: $\mathbf{P}(X|y) = \langle 0.1, 0.9 \rangle$

Key Ideas

Probability Measures Belief

$P(\mathit{false}) = 0 \quad P(\mathit{true}) = 1 \quad \sum_x P(X = x) = 1$

Types of probabilities

- Prior, $P(X = x)$
- Joint, $P(X = x, Y = y)$
- Conditional, $P(X = x | Y = y)$

Inference

- Sums over full joint distribution
- Conditional independence + product and Bayes' rule