

# CSE 150A 250A

# AI: Probabilistic Methods

Fall 2025

Instructor: Trevor Bonjour

*Slides adapted from previous versions of the course (Prof. Lawrence, Prof. Alvarado, Prof Berg-Kirkpatrick)*

# Probability Review

# Probability in AI

Probability Theory == "How knowledge affects belief" (Poole and Mackworth)

What is the  
probability that it is  
raining out?



How much do I  
believe that it is  
raining out?

Viewing probability as measuring belief (rather than frequency of events) is known as the **Bayesian view** of probability (as opposed to the **frequentist view**).

# Discrete Random Variables

Discrete random variables, denoted with capital letters: e.g.,  $X$

Domain of possible values for a variable, denoted with lowercase letters: e.g.,  $\{x_1, x_2, x_3, \dots, x_n\}$

Example: Weather  $W$  ;  $\{w_1 = \text{sunny}, w_2 = \text{cloudy}\}$

# Unconditional (prior) Probability

marginal

$$P(X = x)$$

e.g., What is the probability that the weather is sunny?

$$P(\underline{W} = \underline{w_1})$$

# Axioms of Probability

$$P(X = x) \geq 0$$

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

$$P(X = x_i \text{ or } X = x_j) = P(X = x_i) + P(X = x_j) \text{ if } x_i \neq x_j$$

**Mutually Exclusive!**

# Conditional Probability

$$P(X = x_i | Y = y_j)$$

*Handwritten red annotations: an arrow points from the word "given" to the vertical bar, and the variable  $x_i$  is circled in red.*

"What is my belief that  $X = x_i$  if I **already know**  $Y = y_j$ "

Sometimes, knowing  $Y$  gives you information about  $X$ , i.e., **changes your belief** in  $X$ . In this case  $X$  and  $Y$  are said to be **dependent**.

$$P(X = x_i | Y = y_j) \neq P(X = x_i)$$

*Handwritten red annotations: the inequality symbol  $\neq$  is underlined in red.*

# Webclicker

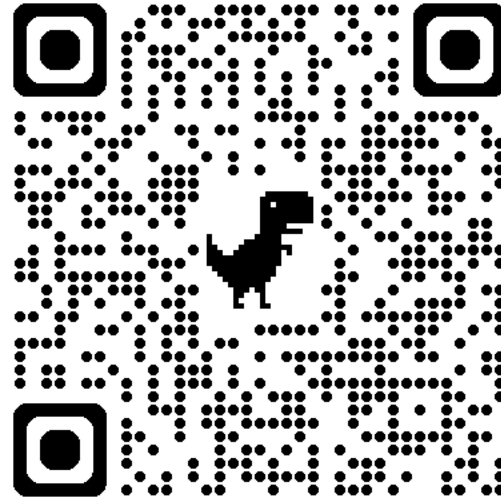
Link:

<https://webclicker.web.app/>

Login using UCSD  
Google account

If NO UCSD account  
—

Use a personal  
Google account



Course code:  
**KSALDG**





# Axioms of Conditional Probability

Which of the following axioms hold for conditional probabilities?

A.  $P(X = x_i | Y = y_j) \geq 0$

B.  $\sum_i P(X = x_i | Y = y_j) = 1$

C.  $\sum_j P(X = x_i | Y = y_j) = 1$  ✗

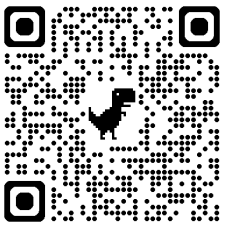
$\sum_i P(X = x_i) = 1$

$1 \underline{\hspace{1cm}} \underline{\hspace{1cm}} \underline{\hspace{1cm}}$

D. *A and B only*

E. *A, B and C*

# Marginal Independence



Course code: KSALDG

$$P(X = \underline{x_i} | Y = y_j) = P(\underline{X = x_i})$$

Sometimes knowing Y **does not change** your belief in X. In this case, X and Y are said to be *independent*.

$$P(W = w_i | Y = y_j) = P(W = w_i)$$

If  $W$  denotes the weather today, For which variable  $Y$  is the above statement most likely true?

- A.  $Y$  = The weather yesterday ✗
- ✓ B.  $Y$  = The day (Mon, Tue...) of the week ✗
- C.  $Y$  = The temperature



Course code: **KSALDG**

# More independence

Consider two students Roberto and Sabrina, who both took the **same** test. Define the following random variables:

$R$  = Roberto aced the test

$S$  = Sabrina aced the test

*\*Assume both students have **similar** ability and the only deciding factor for acing the test is the **test difficulty**.*

What is the most logical relationship between  $P(R = 1)$  and  $P(R = 1|S = 1)$ ?

A.  $P(R = 1) = P(R = 1|S = 1)$

B.  $P(R = 1) > P(R = 1|S = 1)$

C.  $P(R = 1) < P(R = 1|S = 1)$



# Conditional Independence

What if you also know the test was easy (variable T)?

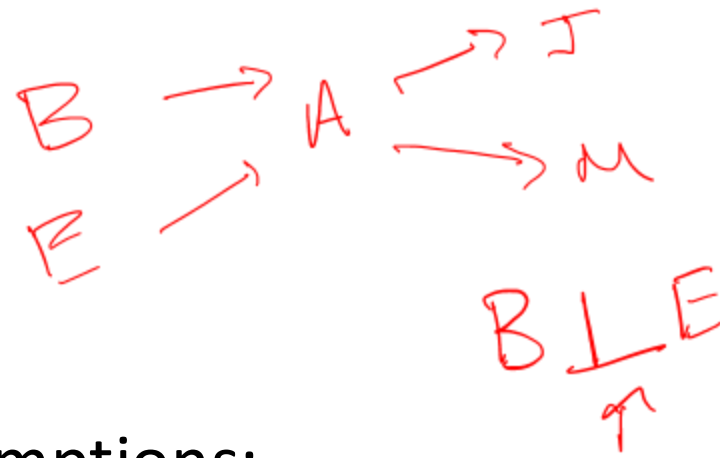
A.  $P(R = 1|T = 1) = P(R = 1|T = 1, S = 1)$

B.  $P(R = 1|T = 1) > P(R = 1|T = 1, S = 1)$

C.  $P(R = 1|T = 1) < P(R = 1|T = 1, S = 1)$

R and S are **conditionally independent** given T. i.e., if you already know T, knowing S does not give you additional information about R.

# More independence



Consider these random variables:

- **B** = Burglary occurs
- **E** = Earthquake occurs
- **A** = Alarm goes off
- **J** = Jamal calls
- **M** = Maya calls

Assumptions:

- Burglary and earthquake are independent. ( $B \perp E$ )
- Alarm can be triggered by either burglary or earthquake.
- Jamal and Maya call if they hear the alarm.
- Jamal and Maya do not directly influence each other.

# Cumulative Evidence



If you know there was an earthquake, what happens to your belief about the alarm going off? Does it increase, decrease or stay the same?

$$P(A = 1) \text{ } \underline{\text{<}} \text{ } P(A = 1|E = 1)$$

If you know that there was an earthquake **AND** that there was a burglary. What happens to your belief about the alarm going off?

$$P(A = 1|E = 1) \text{ } \underline{\text{<}} \text{ } P(A = 1|E = 1, B = 1)$$

# Explaining Away



If the alarm goes off, what happens to your belief about the burglary?

$$P(B = 1) \text{ \textcolor{red}{<} } P(B = 1|A = 1)$$

Now, if the alarm goes off, and you also get to know there was an earthquake – what happens to your belief about the burglary?

$$P(B = 1|A = 1) \text{ \textcolor{red}{>} } P(B = 1|A = 1, E = 1)$$

# Conditional Independence



- If you **don't know** the alarm status, does Jamal's call change your belief about Maya calling?

$$P(M = 1) \leq P(M = 1 | J = 1)$$

- If you know the alarm went off, does Jamal's call still change your belief about Maya calling?

$$P(M = 1 | A = 1) = P(M = 1 | A = 1, J = 1)$$



# Independence



Course code: **GJLOWD**

$$P(X = x_i | Y = y_j) \neq P(X = x_i) \Rightarrow \\ P(X = x_i | Y = y_j, Z = z_k) \neq P(X = x_i | Z = z_k) ?$$

If X and Y are dependent, does this imply X and Y are dependent given Z?

A. Yes

B. No

# Independence



Course code: **GJLOWD**

$$P(X = x_i | Y = y_j) = P(X = x_i) \Rightarrow \\ P(X = x_i | Y = y_j, Z = z_k) = P(X = x_i | Z = z_k) ?$$

If X and Y are independent, does this imply X is independent of Y given Z?

A. Yes

B. No

# Review - Independence

- Marginal Independence

- $P(X = x_i | Y = y_j) = P(X = x_i)$

- Conditional Independence

- $P(X = x_i | Y = y_j) \neq P(X = x_i)$

- $P(X = x_i | Y = y_j, Z = z_k) = P(X = x_i | Z = z_k)$

- Conditional Dependence

- $P(X = x_i | Y = y_j) = P(X = x_i)$

- $P(X = x_i | Y = y_j, Z = z_k) \neq P(X = x_i | Z = z_k)$

# Joint Probability

$$P(X = x_i, Y = y_j)$$

"What is my belief that  $X = x_i$  **and that**  $Y = y_j$ "



Course code: GJLOWD

# Joint Probability

*Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice and participated in anti-nuclear demonstrations.*

Which is more probable?

- A. Linda is a bank teller. 0.001
- B. Linda is a bank teller and is active in the feminist movement.





Course code: GJLOWD

# Joint Probability

$$P(X = x_i, Y = y_j)$$

"What is my belief that  $X = x_i$  **and that**  $Y = y_j$ "

Which of the following is always true?

A.  $P(X = x_i \text{ or } Y = y_j) \leq P(X = x_i, Y = y_j)$

**B.**  $P(X = x_i \text{ or } Y = y_j) \geq P(X = x_i, Y = y_j)$

C.  $P(X = x_i \text{ or } Y = y_j) = P(X = x_i, Y = y_j)$

D. *None of the above.*

# Joint Probability

$$P(X = x_i, Y = y_j)$$

"What is my belief that  $X = x_i$  **and that**  $Y = y_j$ "

$$T \in \{t_1 = \text{hot}, t_2 = \text{cool}\}$$

$$W \in \{w_1 = \text{sunny}, w_2 = \text{cloudy}\}$$

Joint Probability Table

	T = <u>hot</u>	T = cool
W = <u>sunny</u>	<u>0.5</u>	0.3
W = cloudy	0.05	0.15

# Marginalization

$$P(X = x_i) = \sum_j P(X = x_i, Y = y_j)$$

$T \in \{t_1 = \text{hot}, t_2 = \text{cool}\}$

$W \in \{w_1 = \text{sunny}, w_2 = \text{cloudy}\}$

$$\sum_j P(W = w_i, T = t_j)$$

*Handwritten notes:  $w_i = \text{sunny}$ ,  $t_j = \text{hot}$*

Joint Probability Table

	T = hot	T = cool
W = sunny	0.5	0.3
W = cloudy	0.05	0.15

What is  $P(W=...)$ ?

W = sunny	<u>          </u>
W = cloudy	<u>          </u>



# Product Rule

Connects **joint** probability with **conditional** probability

$$P(X = x_i, Y = y_j) = P(X = x_i) P(Y = y_j | X = x_i)$$

$$P(X = x_i, Y = y_j) = P(Y = y_j) P(X = x_i | Y = y_j)$$

# Product Rule Generalization

$$\begin{aligned} P(X = x_i, Y = y_j, Z = z_k, \dots) \\ = P(X = x_i)P(Y = y_j|X = x_i)P(Z = z_k|X = x_i, Y = y_j) \dots \end{aligned}$$

This decomposition often makes reasoning easier.

# Shorthand

Implied Universality:

$$P(X, Y) = P(X|Y)P(Y) = P(Y|X)P(X)$$

Implied Assignment:

$$P(x, y, z) = P(X = x, Y = y, Z = z)$$

# Bayes' Rule

$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)}$$

Derive Bayes' Rule from the Product Rule.

Handwritten derivation of Bayes' Rule from the Product Rule:

$$P(X, Y) = P(X) P(Y|X)$$
$$P(Y) P(X|Y) = P(X) P(Y|X)$$

The derivation shows the joint probability  $P(X, Y)$  expressed as  $P(X)P(Y|X)$  and then as  $P(Y)P(X|Y)$ . A horizontal arrow points from the  $P(X)P(Y|X)$  term in the second equation to the  $P(X)P(Y|X)$  term in the first equation. A curved arrow points from the  $P(X|Y)$  term in the second equation to the  $P(X|Y)$  term in the final Bayes' Rule formula.

# Bayes' Rule example

$$\frac{P(T=1|C=1) \cdot P(C=1)}{\rightarrow P(T=1) \neq}$$

- Let's assume breast cancer affects 1% of all patients  $P(C=1) = 0.01$
- When a patient has cancer, a mammogram test comes back positive 90% of the time (true positive)  $P(T=1|C=1)$
- When a patient does not have cancer, a mammogram test comes back positive 10% of the time (false positive)  $P(T=1|C=0)$

A patient has a positive mammogram test. What is the probability the patient has cancer?  $P(C=1|T=1)$

Define the following random variables:

$\rightarrow T \rightarrow$  Test is positive  $T=1$   
 $\rightarrow C \rightarrow$  Patient has cancer  $C=1$

$T=0$   $\times$   
 $C=0$   $\times$

# Bayes' Rule example

- Let's assume breast cancer affects 1% of all patients ->  $P(C=1) = 0.01$
- When a patient has cancer, a mammogram test comes back positive 90% of the time (true positive) ->  $P(T=1 | C=1) = 0.9$
- When a patient does not have cancer, a mammogram test comes back positive 10% of the time (false positive) ->  $P(T=1 | C=0) = 0.1$

A patient has a positive mammogram test. What is the probability the patient has cancer?

Define the following random variables:

T -> Test is positive

C -> Patient has cancer

Given:  $P(C=1)$ ,  $P(T=1|C=1)$ ,  $P(T=1|C=0)$

What is  $P(C = 1|T = 1)$ ?

$$P(C = 1|T = 1) = \frac{P(T=1|C=1)P(C=1)}{P(T=1)}$$

Bayes' Rule

What is the  $P(T = 1)$ ?

$$P(T = 1) = \sum_c P(T = 1, C = c)$$

Marginalization

$$P(T = 1) = \sum_c P(T = 1|C = c)P(C = c)$$

Product Rule

$$P(T = 1) = P(T = 1|C = 1)P(C = 1) + P(T = 1, C = 0)P(C = 0)$$

What is the  $P(C = 0)$ ?

$$P(C = 0) = 1 - P(C = 1)$$

Axiom

# Conditioning on Background Evidence

- Product rule without background evidence:  $P(X, Y) = P(Y|X)P(X)$
- Claim:  $P(X, Y|E) = P(Y|X, E)P(X|E)$

Prove the conditional version of the product rule. (HW)



# Conditioning on Background Evidence

- Product rule (conditional):  $P(X, Y|E) = P(Y|X, E)P(X|E)$

- Bayes' rule (conditional):

$$P(X|Y, E) = \frac{P(Y|X, E)P(X|E)}{P(Y|E)}$$

- Marginalization (conditional):

$$P(X|E) = \sum_y P(X, Y = y|E)$$

That's all folks!