# CSC421 Intro to Artificial Intelligence

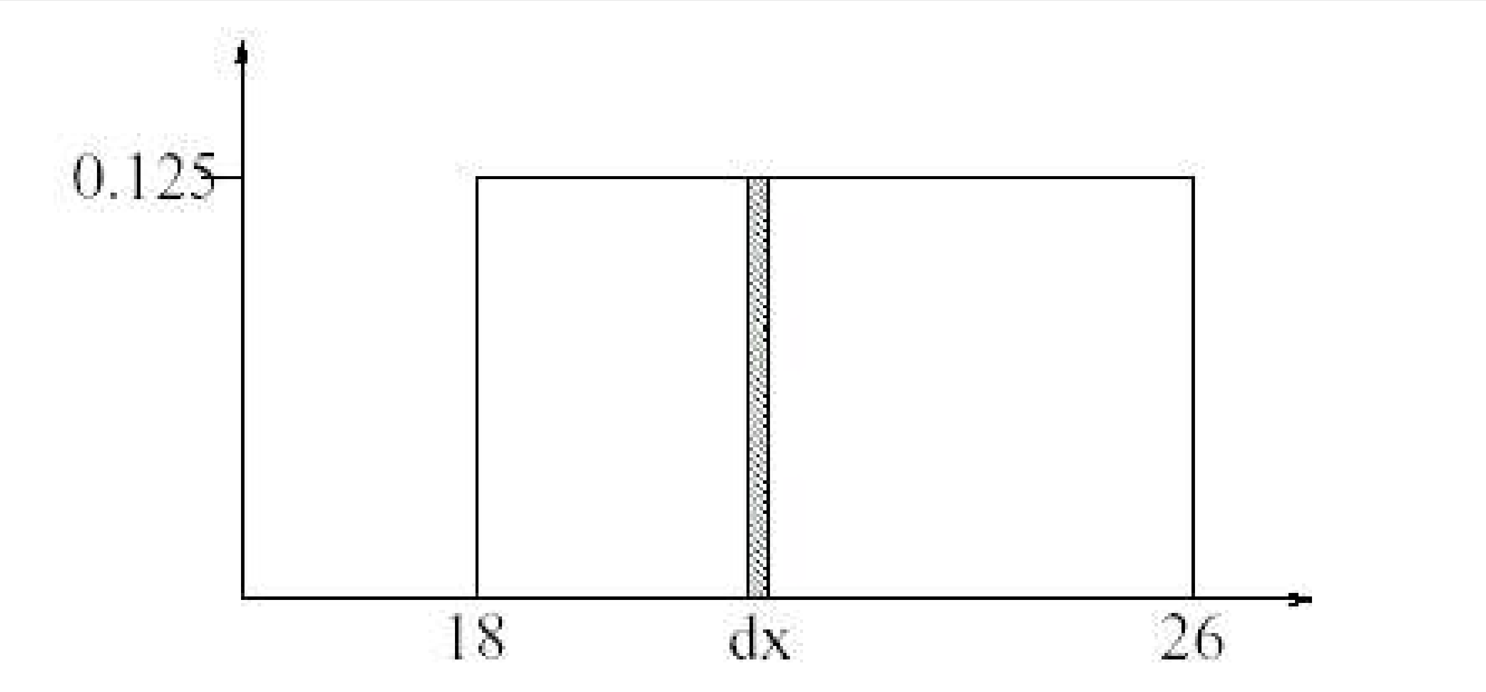
UNIT 21: Uncertainty

# Midterm Review

- Rooks, bishop, pawns
- Edit distance

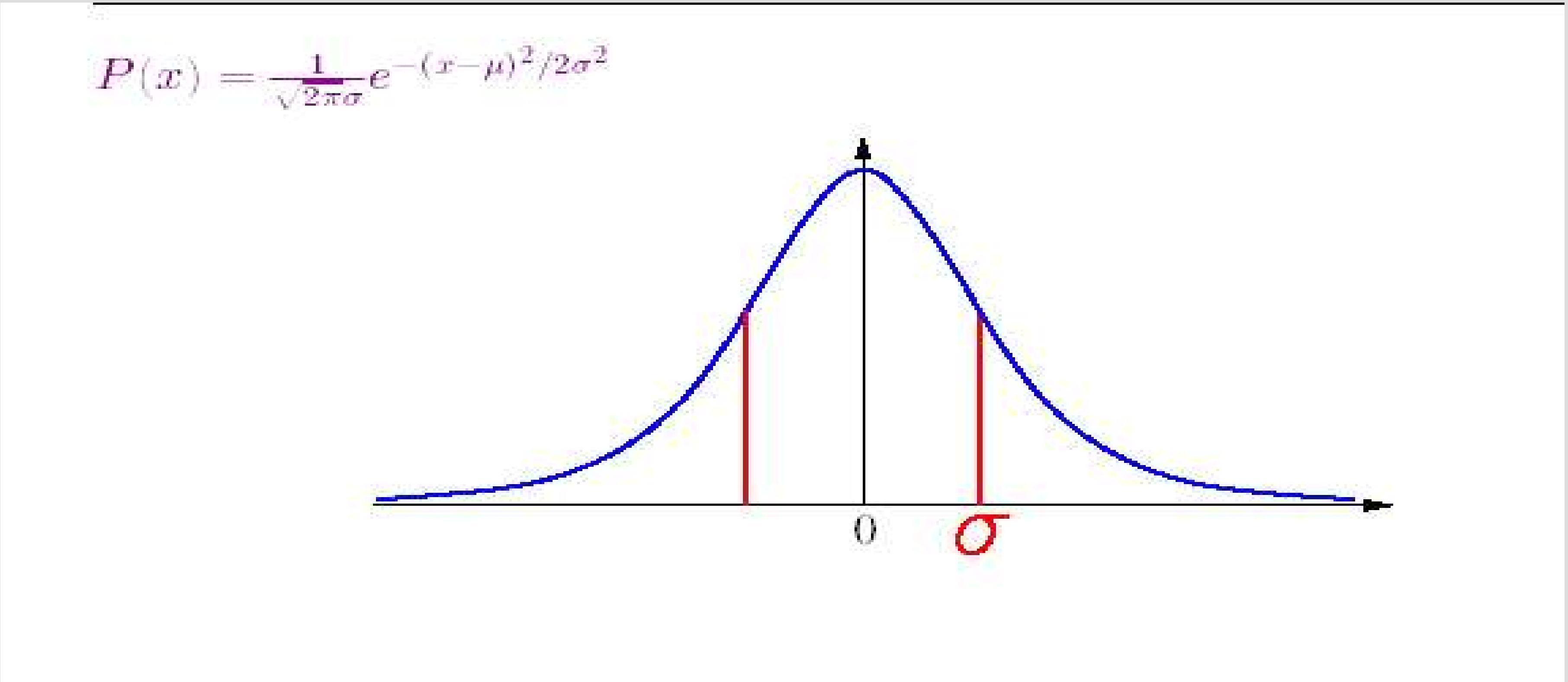
# Probabilities for continuous variables

- Express distribution as a parametrized function of value: $P(X = x) = U[18, 26](x) =$ uniform density between 18, 26
- P density (integrates to 1 )



# Continuous rvs

- P(X = 20.5) = 0.125 really means
  - $\lim_{dx\to 0} P(20.5 \le X \le 20.5 + dx) / dx = 0.125$
- Gaussian Density



# Joint Probability

| Lecture = | boring | ok | interesting |
|---|---|---|---|
| Student = true | 0.144 | 0.04 | 0.016 |
| Student = false | 0.576 | 0.16 | 0.064 |

- Every question about a domain can be answered by the joint probability because every event is a set of sample points (probability is the sum)
- Inference using full joint distributions
  - Marginalization
  - Conditioning

# Conditional probability

- Conditional or posterior probabilities
  - e.g P(student|boring) = 0.8
  - i.e. Given that boring is all I know
  - Not "If Boring = true then 0.8% chance of Student = true. Why ?
- Notation for conditional distribution
  - P(Student | Lecture) = 2-element vector of 2-element vectors
- Product Rule:
  - $P(a \wedge b) = P(b \mid a)\ P(b)$ , when $P(b) > 0$

# Conditional Probability

- If we know more then we have
  - P(student| boring, student) = 1
- New evidence may be irrelevant allowing simplification
  - P(student | boring, raining) = P(student|boring) = 0.8
- This kind of inference, sanctioned by domain knowledge is crucial

# Joint Probability

- Lecture = boring ok interesting
- Student = true 0.144 0.04 0.016
- Student = false 0.576 0.16 0.064
-
- $P(\text{Lecture} = \text{boring} \mid \text{Student} = \text{true}) = ?$

# Independence

- A and B are independent iff
- $P(A \mid B) = P(A)$ or $P(B|A) = P(B)$ or $P(A,B) = P(A)P(B)$
- For N independent biased coins:
  - $2^N \rightarrow N$
- Absolute independence is powerful but rare

# Conditional Independence

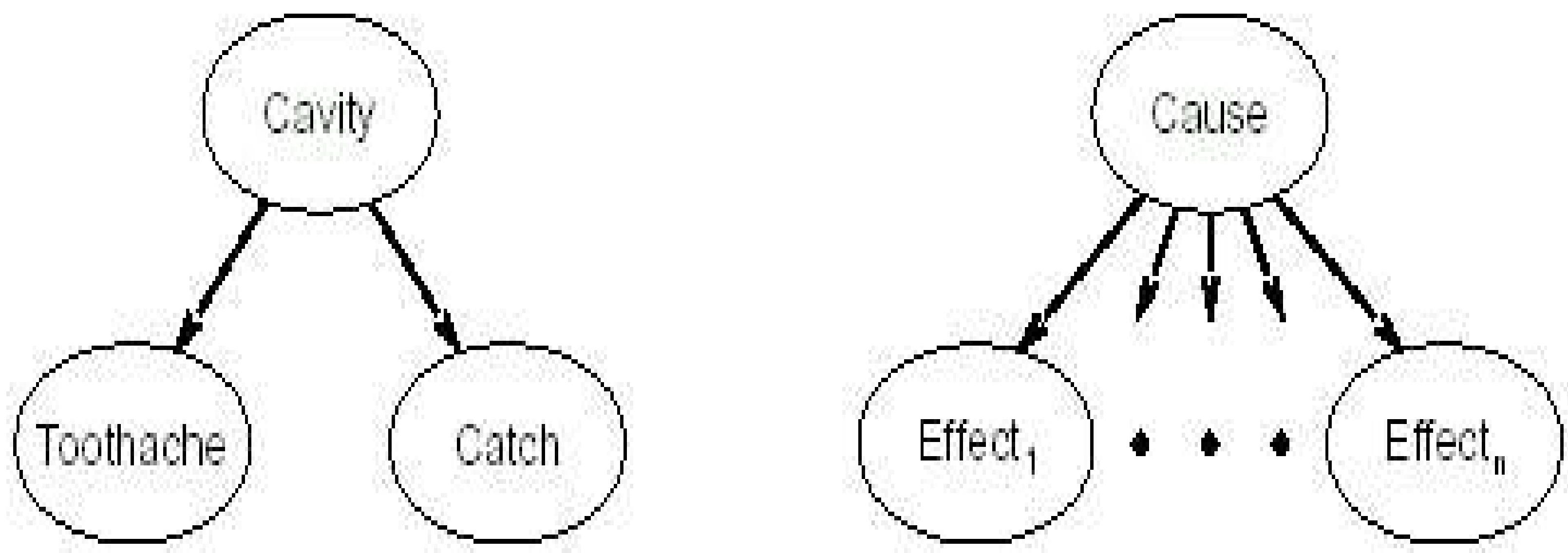
- P(catch | toothache, cavity) = P(catch| cavity)
- P(catch | toothache, not cavity) = P(catch| cavity)
- Catch is conditionally independent of Tootchache given Cavity
- P(Catch | Toothache, Cavity) = P(Catch| Cavity)
- In most cases conditional independence reduces exponential to linear
- MOST BASIC AND ROBUST FORM OF KNOWLEDGE ABOUT UNCERTAINTY

# Bays Rule

- $P(a|b) = P(b|a)\ P(a)\ /\ P(b)$
- How can you prove it ?
- In distribution form
  - $P(Y|X) = P(X|Y)\ P(Y)\ |\ P(X)$
  - Diagnostic probability from causal probability
  - P(Cause |Effect) = P(Effect|Cause) P(Cause) / P(Effect)
  - M menigitis, S stiff neck with:
    - $P(M) = 0.0001$, $P(S|M) = 0.8$ $P(S) = 0.1$
  - $P(m|s) = ?$
- Why is this useful ?
  - Are diagnostic rules or causal rules harder to estimate ?

# Naïve Bayes Model

- Commonly occuring pattern in which a single cause directly influences a number of effects, all of which are conditionally independent



# Naïve Bayes Model

- P(Cause, Effect1, ... EffectN) =
  $P(Cause) \prod P(Effect_i | Cause)$
- Common simplifying assumption in machine learning, data mining
- Bayesian Classifier (not exactly correct from a strict baysian point of view)

# Probabilistic Wumpus World

- Pits cause breezes in neighbors
- Each square (except [1,1] ) contains a pit with probability 0.2
- Random variables
  - P[i,j] for each square
  - B[i,j] for each square
- Please work through examples from textbook