

Think Bayesian



A man is running. Why?

Principle 1: Use prior knowledge

1



He is in a hurry

2



He is doing sports

3



He always runs

4



~~He saw a dragon~~

Low prior probability



A man is running. Why?

Principle 2: Choose answer that explains observations the most

Suppose, for the sake of argument, that he is not wearing a sports suit. Then, 'he is doing sports' doesn't really explain our observation.

1



He is in a hurry

3



He always runs

2



~~He is doing sports
Contradicts the data~~

4



~~He saw a dragon~~



A man is running. Why?

Principle 3: Avoid making extra assumptions

1



He is in a hurry

2



He is doing sports

3



He always runs

Too many assumptions

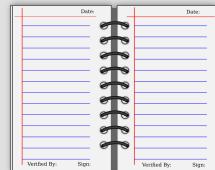
4



He saw a dragon



Main principles



Principle 1:
Use prior knowledge

Principle 2:
Choose answer that explains observations the most

Principle 3:
Avoid making extra assumptions

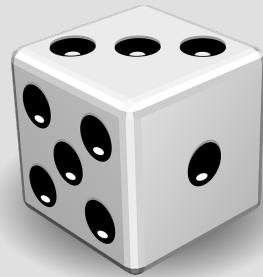
Principle 3 is also known as Occam's razor.



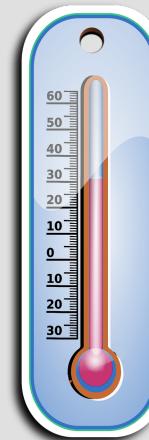
Review of probability



Random variables



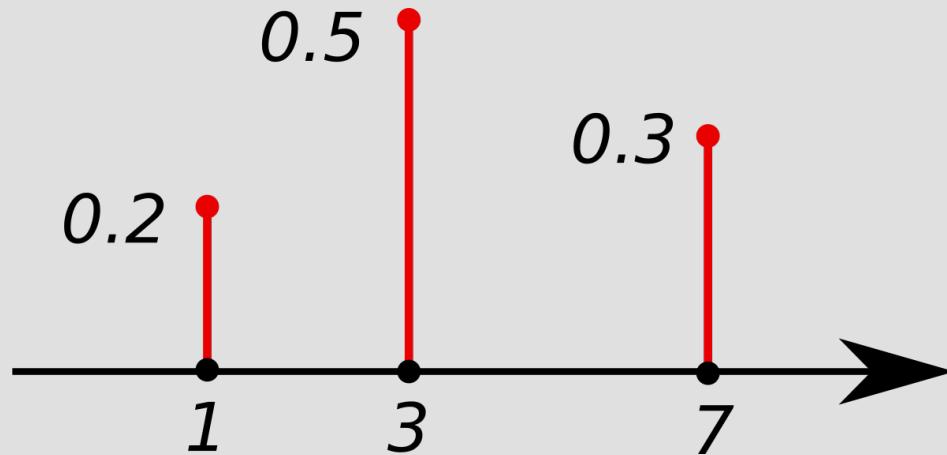
Discrete



Continuous



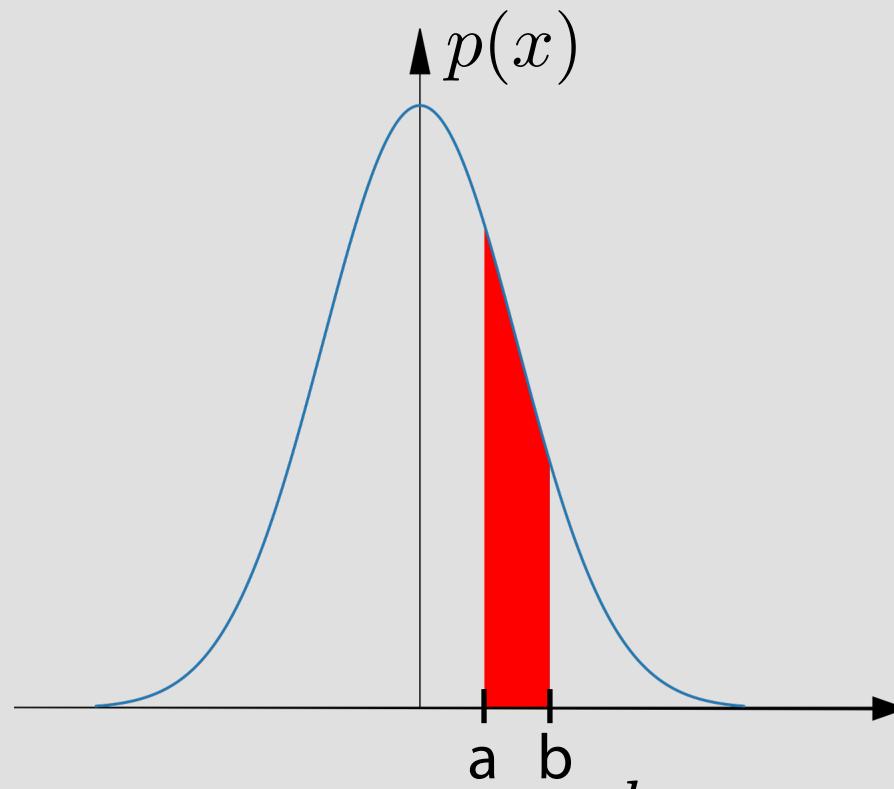
Discrete: Probability Mass Function (PMF)



$$P(X) = \begin{cases} 0.2 & X = 1 \\ 0.5 & X = 3 \\ 0.3 & X = 7 \\ 0 & otherwise \end{cases}$$



Continuous: Probability Density Function (PDF)



$$P(x \in [a, b]) = \int_a^b p(x) dx$$



Independence

X and Y are independent if:

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

↑
Joint Marginals
↓ ↓



Conditional probability

Probability of **X** given that **Y** happened:

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

Joint

Conditional

Marginal

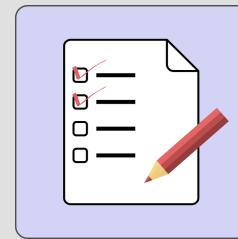
The diagram illustrates the formula for conditional probability. It features a central fraction $P(X, Y) / P(Y)$. A red bracket on the left side of the numerator $P(X, Y)$ is labeled "Conditional". A red bracket on the right side of the denominator $P(Y)$ is labeled "Marginal". Above the fraction, a red bracket spanning both the numerator and the denominator is labeled "Joint".



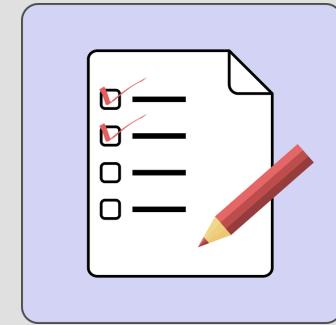
Conditional probability

$$P(M) = 0.4$$

$$P(M \& F) = 0.25$$



Midterm



Final

$$P(F|M) = \frac{P(M \& F)}{P(M)} = \frac{0.25}{0.4} = 0.625$$



Chain rule



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

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

$$P(X_1, \dots, X_N) = \prod_{i=1}^N P(X_i | X_1, \dots, X_{i-1})$$



Sum rule

Marginalization

$$p(X) = \int_{-\infty}^{\infty} p(X, Y) dY$$



Bayes theorem

θ — parameters

This is important!

X — observations

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

Evidence

The diagram illustrates the components of Bayes' theorem. At the top, four labels are positioned: 'Posterior' (left), 'Likelihood' (center), 'Prior' (right), and 'Evidence' (bottom). Red arrows point from each label to its corresponding term in the equation below. The 'Posterior' arrow points to the first term in the numerator. The 'Likelihood' arrow points to the first term in the fraction. The 'Prior' arrow points to the second term in the numerator. The 'Evidence' arrow points to the denominator.

