Welcome to CS109A

Gili Rusak

Agenda

- Conditional probability*
    - Bayes' Theorem*
    - Independence*
- Random variables (more next week)
    - Probability mass function
    - Cumulative distribution function
    - Expectation*
- Applications

* Relevant for HW2

Conditional Probability

Definitions

Definition of Conditional Probability

The probability of E given that (aka conditioned on) event F already happened:

$$P(E \mid F) = \frac{P(EF)}{P(F)} = \frac{P(E \cap F)}{P(F)}$$

The Law of Total Probability

For events E and F,

$$P(F) = P(F \mid E)P(E) + P(F \mid E^C)P(E^C)$$

Bayes' Theorem

Bayes' Theorem
The most common form of Bayes' Theorem is:

$$P(E \mid F) = \frac{P(F \mid E)P(E)}{P(F)}$$

Independence

Independence
Two events, E and F, are **independent** if and only if:

$$P(EF) = P(E)P(F)$$

Otherwise, they are called **dependent** events.

Two events E and F are called **conditionally independent** given a third event G, if

$$P(EF \mid G) = P(E \mid G)P(F \mid G)$$

Or, equivalently:

$$P(E \mid FG) = P(E \mid G)$$

Example Problem 1.1

Consider a scenario in which you have three coins. Coin A comes up heads with probability 0.5, Coin B with probability 0.2, and Coin C with probability 0.9. You choose a coin randomly with equal probability for all three.

Problem: What is the probability that you choose Coin A, i.e. $P(A)$?

Hint: You can read off the answer from the problem statement! No calculation is needed.

* Problem by Alex Tsun

Example Problem 1.1

Consider a scenario in which you have three coins. Coin A comes up heads with probability 0.5, Coin B with probability 0.2, and Coin C with probability 0.9. **You choose a coin randomly with equal probability for all three.**

What is the probability that you choose Coin A, i.e. P(A)?

Hint: You can read off the answer from the problem statement! No calculation is needed.

Solution: 0.3333

Example Problem 1.2

Consider a scenario in which you have three coins. Coin A comes up heads with probability 0.5, Coin B with probability 0.2, and Coin C with probability 0.9. You choose a coin randomly with equal probability for all three.

Problem: Given that you choose Coin A, what is the probability you get heads, i.e. $P(H|A)$?

Hint: You can read off the answer from the problem statement! No calculation is needed.

Example Problem 1.2

Consider a scenario in which you have three coins. Coin A comes up heads with probability 0.5, Coin B with probability 0.2, and Coin C with probability 0.9. You choose a coin randomly with equal probability for all three.

Problem: Given that you choose Coin A, what is the probability you get heads, i.e. $P(H|A)$?

Hint: You can read off the answer from the problem statement! No calculation is needed.

Solution: 0.5

Example Problem 1.3

Consider a scenario in which you have three coins. Coin A comes up heads with probability 0.5, Coin B with probability 0.2, and Coin C with probability 0.9. You choose a coin randomly with equal probability for all three.

Problem: Use one of the definitions from the slides to calculate the probability you get heads, i.e. P(H). Hint: Multiple cases.

Example Problem 1.3

Key idea: law of total probability

Consider a scenario in which you have three coins. Coin A comes up heads with probability 0.5, Coin B with probability 0.2, and Coin C with probability 0.9. You choose a coin randomly with equal probability for all three.

Problem: Use one of the definitions from the slides to calculate the probability you get heads, i.e. P(H). Hint: Multiple cases.

Solution: By Law of total probability

$P(H) = P(H \mid A) * P(A) + P(H|B) * P(B) + P(H \mid C) * P(C)$

$= 0.5 * \frac{1}{3} + 0.2 * \frac{1}{3} + 0.9 * \frac{1}{3} = 1.6/3 =$ **0.5333**

Example Problem 1.4

Consider a scenario in which you have three coins. Coin A comes up heads with probability 0.5, Coin B with probability 0.2, and Coin C with probability 0.9. You choose a coin randomly with equal probability for all three.

Problem: Use one of the definitions from the slides to calculate the probability that you chose the Coin A, given that you got heads, i.e. P(A|H).

Example Problem 1.4 Solution

Key idea: Bayes' Theorem

Consider a scenario in which you have three coins. Coin A comes up heads with probability 0.5, Coin B with probability 0.2, and Coin C with probability 0.9. You choose a coin randomly with equal probability for all three.

Problem: Use one of the definitions from the slides to calculate the probability that you chose the Coin A, given that you got heads, i.e. P(A|H).

Solution: By Bayes Theorem,

$P(A|H) = P(H \mid A) * P(A) / P(H)$

$= 0.5 * \frac{1}{3} / (0.5 * \frac{1}{3} + 0.2 * \frac{1}{3} + 0.9 * \frac{1}{3})$

$= 0.5 * \frac{1}{3} / 0.533 = (0.3125)$

Example Problem 2 [Spring 2019 HW2]

A robot can be in exactly **one of two locations: L1 or L2**. The probability that the robot is in location L1 is P(L1) and the probability it is in location L2 is P(L2). Based on all past observations, the robot thinks that there is a **0.8 probability it is in L1** and a **0.2 probability that it is in L2**.

The robot's vision algorithm detects a window, and although there is only a window in L2, it can't conclude that it is in fact in L2 because its image recognition algorithm is not perfect.

The probability of observing a window given there is **no** window at its location is 0.2 and the probability of observing a window given there **is** a window is 0.9. After incorporating the observation of a window, what is the robot's new values for P(L1) and P(L2)?

Example Problem 2 Solution

Key idea: translate paragraph to mathematical notation

Let O be the event that we observe a window.

Problem Goal: "After incorporating the observation of a window, what is the robot’s new values for P(L1) and P(L2)?"

Example Problem 2 Solution

Key idea: translate paragraph to mathematical notation

Let O be the event that we observe a window.

Problem Goal: "After incorporating the observation of a window, what is the robot’s new values for P(L1) and P(L2)?"

i.e. calculate P(L1 | O) and P(L2 | O).

Example Problem 2 Solution

Key idea: translate paragraph to mathematical notation

Let O be the event that we observe a window.

Problem Goal: Calculate P(L1 | O) and P(L2 | O).

Key information in the problem:

"A robot can be in exactly **one of two locations: L1 or L2**."

Example Problem 2 Solution

Key idea: translate paragraph to mathematical notation

Let O be the event that we observe a window.

Problem Goal: Calculate P(L1 | O) and P(L2 | O).

Key information in the problem:

"A robot can be in exactly **one of two locations: L1 or L2**."

1. Since L1 and L2 are the only locations, $L1 = L2^C$ and vice versa.

Example Problem 2 Solution

Key idea: translate paragraph to mathematical notation

Let O be the event that we observe a window.

Problem Goal: Calculate P(L1 | O) and P(L2 | O).

Key information in the problem:

"Based on all past observations, the robot thinks that there is a **0.8 probability it is in L1** and a **0.2 probability that it is in L2**."

Example Problem 2 Solution

Key idea: translate paragraph to mathematical notation

Let O be the event that we observe a window.

Problem Goal: Calculate P(L1 | O) and P(L2 | O).

Key information in the problem:

"Based on all past observations, the robot thinks that there is a **0.8 probability it is in L1** and a **0.2 probability that it is in L2**."

2. P(L1) = 0.8 and P(L2) = 0.2.

Example Problem 2 Solution

Key idea: translate paragraph to mathematical notation

Let O be the event that we observe a window.

Problem Goal: Calculate P(L1 | O) and P(L2 | O).

Key information in the problem:

"The probability of observing a window given there is **no** window at its location is 0.2 and the probability of observing a window given there **is** a window is 0.9."

Example Problem 2 Solution

Key idea: translate paragraph to mathematical notation

Let O be the event that we observe a window.

Problem Goal: Calculate P(L1 | O) and P(L2 | O).

Key information in the problem:

"The probability of observing a window given there is **no** window at its location is 0.2 and the probability of observing a window given there **is** a window is 0.9."

3. Therefore, P(O|L1) = 0.2 and P(O|L2) = 0.9.

Example Problem 2 Solution

Key idea: translate paragraph to mathematical notation

Let O be the event that we observe a window.

Problem Goal: Calculate P(L1 | O) and P(L2 | O).

Key information in the problem:

1. Since L1 and L2 are the only locations, $L1 = L2^C$ and vice versa.

2. P(L1) = 0.8 and P(L2) = 0.2.

3. Therefore, P(O|L1) = 0.2 and P(O|L2) = 0.9.

Example Problem 2 Solution

Key idea: Bayes' Theorem

Use Bayes' theorem.

$$P(L_1 \mid O) = \frac{P(O \mid L_1)P(L_1)}{P(O \mid L_1)P(L_1) + P(O \mid L_1^C)P(L_1{}^C)}$$

$$= \frac{P(O \mid L_1)\ (0.8)}{P(O \mid L_1)\ (0.8) + P(O \mid L_1^C)\ (0.2)}$$

= (0.2) * (0.8) / (0.2 * 0.8 + 0.9 * 0.2)

Random Variables

Random Variables

- A **random variable** (RV) is a variable that probabilistically takes on different values.
- The **probability mass function** (PMF) of a random variable is a function that maps possible outcomes of a random variable to the corresponding probabilities.

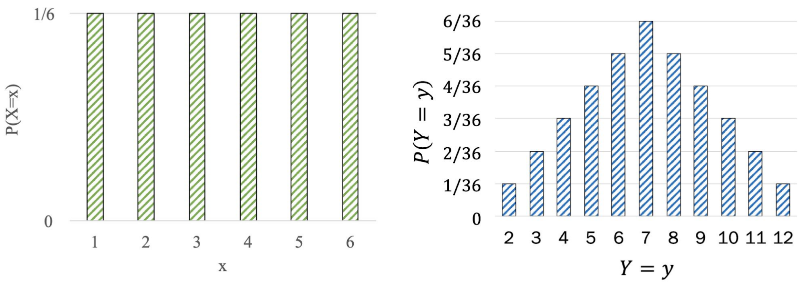

Figure 1: On the left, the PMF of a single 6 sided die roll. On the right, the PMF of the sum of two dice rolls.

- The **cumulative distribution function** (CDF) of a random variable X is a function F specified as $F(a) = P(X \leq a)$, the probability that X takes on a value less than or equal to some value a.

Expectation

The **expectation** of a discrete random variable X is defined as:

$$E[X] = \sum_{x:P(x)>0} xP(x)$$

Preview for week 3 and 4 in class

- Common random variables
  - Bernoulli RV
  - Binomial RV
  - Geometric RV
  - Poisson RV
  - Normal RV
  - Others!
- Properties of RVs
  - PDF, CDF, E[X], applications

Applications

Applications

- Naive bayes' model: a machine learning model (end of course)
- Classification machine learning problems
- Healthcare and biology
  - Punnett squares (see HW2)

See you next Tuesday!

Example Problem 2 [Midterm Spring 2020]

A home robot has two different sensors for motion detection. If there is a moving object, sensor V (video camera) will detect motion with probability 0.95, and sensor L (laser) will detect motion with probability 0.8. If there is no moving object, there is a 0.1 probability that sensor V will detect motion (even though there is no object), and a 0.05 probability that sensor L will detect motion.

Based on empirical evidence, the probability that there is a moving object is 0.7. Note that these sensors use independent detection algorithms to identify motion, so that **conditioned** on there being a moving object (or not), the events of detecting motion (or not) for each Sensor is **independent**.

Problem: Given that there is a moving object and that sensor V does not detect motion, what is the probability that sensor L detects motion? Give a numerical answer.