CSCI 3022: Intro to Data Science

Lecture 5: Axioms & Theorems of Probability

Rachel Cox

Department of Computer Science

Announcements & Reminders

HW1 Due Friday at 11:59pm

What will we learn today?

- ❑ Conditional Probability
- ❑ Independence
- ❑ Law of Total Probability

- ❑ *A Modern Introduction to Probability and Statistics, Chapter 2 & 3*

Probability – Warm-up

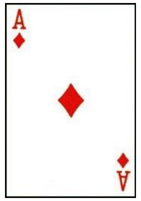
$\cup$ or

$\cap$ and

A **probability function** assigns a value in $[0, 1]$ to each outcome or event in the sample space Ω. Additionally, we know that

- $P(\Omega) = 1$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- Which simplifies to $P(A \cup B) = P(A) + P(B)$ if A and B are disjoint.

Example: Suppose you draw a single card from a standard 52-card deck.

What is the probability that the card is the  ?

What is the probability that the card is an Ace or a Diamond?

$$P(A\Diamond) = \frac{1}{52}$$

$$P(\text{Ace} \cup \Diamond) = P(\text{Ace}) + P(\Diamond) - P(\text{Ace} \cap \Diamond)$$

$$= 4/52 + 13/52 - 1/52 \; \boxed{= 16/52}$$

Probability

We can compute the probability of an event A as the sum of the probabilities of all of the outcomes in A:

$$P(A) = \sum_{\omega \in A} P(\omega)$$

events

Example: Suppose that we flip a biased coin with the probability function $P(\{H, T\}) = \{p, 1-p\}$ three times. What is the probability that we get two or more Tails?

Here: A is the event of getting 2 or more tails.

$A = \{HTT, THT, TTH, TTT\}$

$\omega \in A$

$\underbrace{HTT, THT, TTH, TTT}_{\omega}$

e.g. $p(H) = .1$
$p(T) = .9$

$P(A) = P(HTT) + P(THT) + P(TTH) + P(TTT)$

$P(A) = p(1-p)(1-p) + (1-p)p(1-p) + (1-p)(1-p)p + (1-p)(1-p)(1-p)$

$\boxed{P(A) = 3p(1-p)^2 + (1-p)^3}$

Probability

Note H3 – 3rd flip yield 'Head'

Assume $p > 0$

Example: Now suppose that we flip a biased coin with the probability function $P(\{H, T\}) = \{p, 1-p\}$ *until a Heads comes up*. Show that the probability that we eventually flip a Heads is 1.

disjoint events!

$$P(\text{Eventually flipping } H) = P(\text{Heads on F1} \cup \text{Heads on F2} \cup \text{Heads on F3} \cup \cdots)$$

$$= P(H1) + P(H2) + P(H3) + P(H4) \cdots$$

$$= P(H) + P(TH) + P(TTH) + P(TTTH) + \cdots$$

$$= p + (1-p)p + (1-p)^2 p + (1-p)^3 p + \cdots \text{ to infinity!}$$

$$= p[1 + (1-p) + (1-p)^2 + (1-p)^3 + \cdots]$$

$$= p \sum_{i=0}^{\infty} (1-p)^i$$

This is a geometric series!

$$\boxed{\sum ar^n = \frac{a}{1-r} \quad a = \text{first term}, \ r = \text{common ratio}}$$

Probability

Example: Now suppose that we flip a biased coin with the probability function $P(\{H,T\}) = \{p, 1-p\}$ *until a Heads comes up*. Show that the probability that we eventually flip a Heads is 1.

$$P(\text{Heads Eventually}) = p\sum_{i=0}^{\infty}(1-p)^i$$

is $|1-p|<1$ as long $p \in (0,1)$

$$= p\,\frac{1}{1-(1-p)}$$

$$= \frac{p}{p}$$

$$\boxed{= 1}$$

☺

Probability

Assume uniform distribution of births across all months

Example: You stop a random person on the street and ask them what month they were born in. What is the probability that they were born in a "long" month? ("long" means 31 days)

- Jan, Mar, May, July, Aug, Oct, Dec

$$p(\text{"long" month}) = 7/12$$

What is the probability that they were born in a month with an *r* in the name?

Jan, Feb, Mar, Apr, Sept, Oct, Nov, Dec

$$p(\text{"r" month}) = 8/12$$

Conditional Probability

Example: Now suppose this person tells you that they were born in a long month. What, then, is the probability that they were born in a month with an *r* in the name?

P(person $\underbrace{\text{born in "r" month}}_{R}$ **given** the knowledge they were $\underbrace{\text{born in "long" month}}_{L}$) $= 4/7$

↓given

$$P(R \mid L) = \frac{P(R \cap L)}{P(L)} = \frac{4/12}{7/12} = 4/7$$

Conditional Probability

how many bit strings of length 4:

$\underline{2} \cdot \underline{2} \cdot \underline{2} \cdot \underline{2}$

$= 2^4 = 16$

The conditional probability of A given C is defined by

$$P(A \mid C) = \frac{P(A \cap C)}{P(C)}$$

Example: A bit string (1s and 0s) of length 4 is generated at random so that each of the 16 possible bit-strings is equally likely. What is the probability that it contains at least 2 consecutive 1s, given that the first bit is a 1?

1000
1001
1010
1011
1100
1101
1110
1111

$$\boxed{P(\underbrace{\text{At least 2 consecutive 1's}}_{C} \mid \underbrace{\text{first bit}=1}_{O}) = 5/8}$$

$$P(C \mid O) = \frac{P(C \cap O)}{P(O)} = \frac{5/16}{8/16} \boxed{= 5/8}$$

Conditional Probability

Recall $P(A) + P(A^c) = 1$

$P(A^c) = 1 - P(A)$

Example: A bit string (1s and 0s) of length 4 is generated at random so that each of the 16 possible bit-strings is equally likely. What is the probability that it does not contain 2 consecutive 1s, given that the first bit is a 1?

$$P(C \mid O) = \frac{5}{8}$$

$$P(C^c \mid O) = 1 - P(C \mid O)$$

$$= 1 - 5/8$$

$$\boxed{= 3/8}$$

Conditional Probability

$$P(A|C) = \frac{P(A \cap C)}{P(C)}$$

The following is a product rule (aka multiplication rule) of probability:

$$P(A \cap C) = P(A \mid C)\, P(C)$$

- Useful when the conditional probability is easy to compute, but the probability of the intersections is not.

Example: Suppose you draw two cards from a standard deck. What is the probability that they are both red? Without Replacement.

$$P(R1 \cap R2) = P(R2 \cap R1) = P(R2 \mid R1)\, P(R1)$$

$$= \frac{25}{51} \cdot \frac{26}{52}$$

Probability – Independence

An event A is said to be **independent** of event B if $P(A \mid B) = P(A)$

This definition, combined with the product rule and the definition of conditional probability, give us a few equivalent tests for independence of two events:

1) $P(A \mid B) = P(A)$
2) $P(B \mid A) = P(B)$
3) $P(A \cap B) = P(A)P(B)$

Events $A_1, A_2, \dots, A_m$ are **independent** if $P(A_1 \cap A_2 \cap \dots \cap A_m) = P(A_1)P(A_2) \dots P(A_m)$

Probability – Independence

Example: Suppose you flip a fair coin twice. Let A= "Heads on flip 1", B = "Heads on flip 2", and C = "Same outcome on both flips". Are these events independent or dependent?

$$\Omega = \{HH, HT, TH, TT\}$$

$$p(A \cap B \cap C) \stackrel{?}{=} p(A)p(B)p(C)$$

$$p(A \cap B \cap C) = \frac{1}{4}$$

$$p(A) \cdot p(B) p(C) = \frac{1}{8}$$

$$p(A) = \frac{1}{2} \qquad p(C) = 1/2$$

$$p(B) = 1/2$$

$$\frac{1}{4} \neq \frac{1}{8}$$

$\Rightarrow$ Not independent!

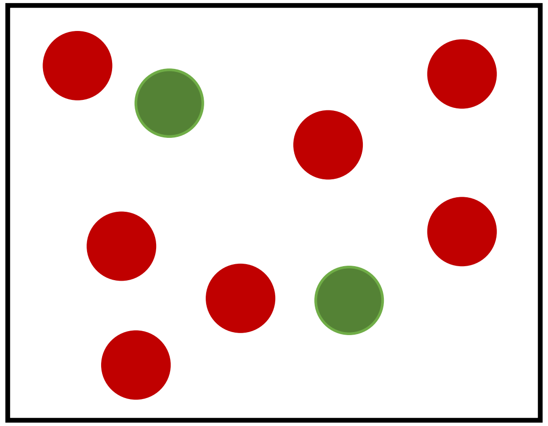
Law of Total Probability

Example: Suppose we have two boxes filled with green and red balls. Paul selects a ball by first choosing one of the two boxes at random. He then selects one of the balls in the box at random. What is the probability Paul has selected a red ball?

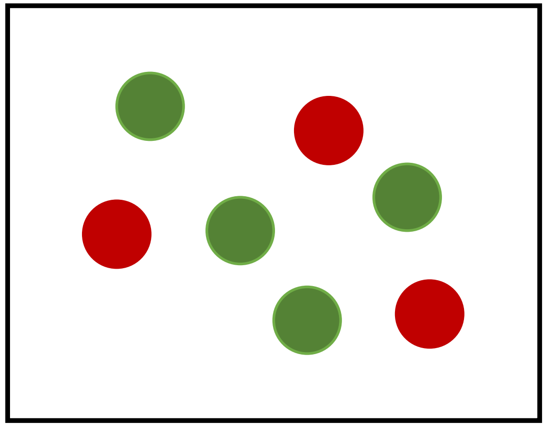
Assume

$P(\text{Box1}) = 1/2$

$P(\text{Box2}) = 1/2$



Box 1: 2 greens, 7 reds



Box 2: 4 greens, 3 reds

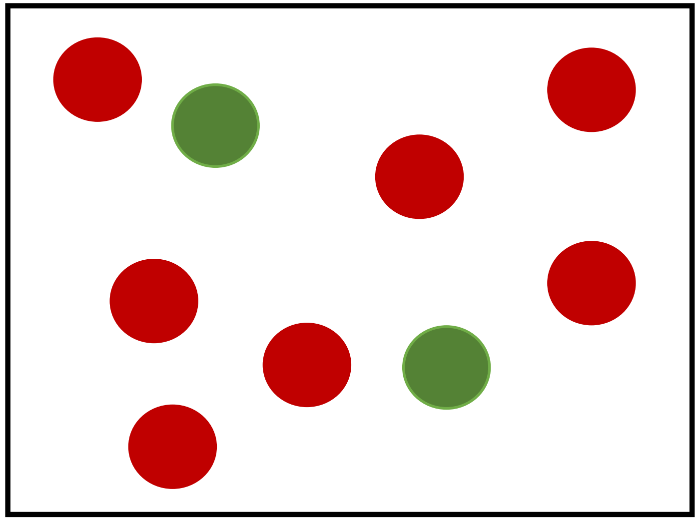
$$P(\text{Red}) = P(\text{Red} \mid \text{Box1})\,P(\text{Box1}) + P(\text{Red} \mid \text{Box2})\,P(\text{Box2})$$

$$= \frac{7}{9} \cdot \frac{1}{2} + \frac{3}{7} \cdot \frac{1}{2}$$

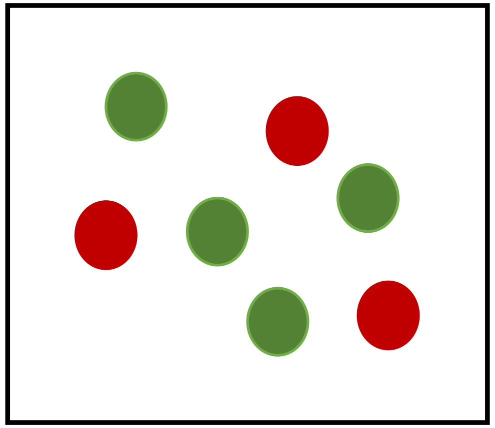
$$\boxed{= \frac{7}{18} + \frac{3}{14}}$$

Law of Total Probability

Example: Same scenario as the previous example, but now suppose that Box 1 is much larger, so that we are twice as likely to grab the first box as the second.



Box 1: 2 greens, 7 reds



Box 2: 4 greens, 3 reds

$$P(\text{Box}1) + P(\text{Box}2) = 1$$

$$P(\text{Box}1) = 2P(\text{Box}2)$$

$$2P(\text{Box}2) + P(\text{Box}2) = 1$$

$$3P(\text{Box}2) = 1$$

$$P(\text{Box}2) = 1/3$$

$$P(\text{Red}) = P(\text{Red} \mid \text{Box}1)P(\text{Box}1) + P(\text{Red} \mid \text{Box}2)P(\text{Box}2)$$

$$= \frac{7}{9}\cdot\frac{2}{3} + \frac{3}{7}\cdot\frac{1}{3}$$

$$\boxed{= 14/27 + 1/7}$$

Law of Total Probability

Suppose $C_1, C_2, \dots, C_m$ are disjoint events such that $C_1 \cup C_2 \cup \dots \cup C_m = \Omega$. Then the probability of an arbitrary event A can be expressed as:

$$P(A) = P(A \mid C_1)\, P(C_1) + P(A \mid C_2)\, P(C_2) + \dots + P(A \mid C_m)\, P(C_m)$$

Next Time:

- Bayes Rule