Discrete Probability Spaces, Conditional Probability, and Independence

CS 27100 Winter 2024
Lecture 17

David Cash
Probability So Far

$\Omega$: sample space (finite)

$E \subseteq \Omega$: event

Probability of $E$ is defined to be

$$\Pr(E) = \frac{|E|}{|\Omega|}$$

→ Only applies when all outcomes equally likely.
This Lecture: More General Probability.

Def A discrete probability space is a pair $(\mathcal{Z}, p)$ where $p: \mathcal{Z} \to \mathbb{R}$ is a function satisfying

1. For all $w \in \mathcal{Z}$, $p(w) > 0$

2. $\sum_{w \in \mathcal{Z}} p(w) = 1$.

The function $p$ is called a distribution.
Probability Measures

Let \((\mathcal{X}, \mathcal{F})\) be a discrete probability space.

\[
\mathcal{F} \subseteq 2^\mathcal{X}
\]

We call the function \(\Pr : 2^\mathcal{X} \to \mathbb{R}\) defined by

\[
\Pr(E) = \sum_{\omega \in E} p(\omega)
\]

the probability measure induced by \(\rho\).

We will often refer to \((\mathcal{X}, \mathcal{F}, \Pr)\) as the discrete probability space (instead of \((\mathcal{X}, \rho)\).
Example 1: Uniform Distribution

Let $\Omega$ be finite and define

$$p(\omega) = \frac{1}{|\Omega|}$$

for all $\omega \in \Omega$. Then for all $E \subseteq \Omega$,

$$P_r(E) = \sum_{\omega \in E} p(\omega) = \sum_{\omega \in E} \frac{1}{|\Omega|} = \frac{|E|}{|\Omega|}.$$

This is called the uniform distribution.
Example 2: Biased Coin

Take $S_2 = \mathcal{H} \cup \mathcal{T}^3$ and define, say,

$p(H) = \frac{3}{4}, \quad p(T) = \frac{1}{4}.$

More generally, we can take any $0 \leq \theta \leq 1$

and define $p(H) = \theta, \quad p(T) = 1 - \theta.$
Example 3: Tossing Until Heads

We toss a fair coin until it shows Heads.

$$S_2 = \emptyset, H, TH, TTH, \ldots \geq 3$$ (infinite!)

$$P(H) = \frac{1}{2}, \quad P(TH) = \frac{1}{4}, \quad P(TTH) = \frac{1}{8}, \ldots$$

$$\Rightarrow P(T^j H) = \left( \frac{1}{2} \right)^{j+1}$$

We check

$$\sum_{w \in S_2} P(w) = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots = 1$$

geometric series
Consequences of the Definition

Theorem: Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a discrete probability space. Then

1. $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\Omega) = 1$

2. $\mathbb{P}(E^c) = 1 - \mathbb{P}(E)$

3. $\mathbb{P}(E \cup F) = \mathbb{P}(E) + \mathbb{P}(F) - \mathbb{P}(E \cap F)$

Proofs are easy generalizations of counting.
Conditional Probability
Conditional Probability: Intuition

"Conditioning" is how to take new information into account.
B (pink): \( \Pr(B) = \frac{6}{12} = \frac{1}{2} \)

A (green): \( \Pr(A) = \frac{3}{12} = \frac{1}{4} \)

Conditional probability of A given B, written \( \Pr(A|B) \) is

\[ \Pr(A|B) = \frac{2}{6} = \frac{1}{3} \]

(Note: \( \Pr(A|B) \neq \Pr(A) \)).
Definition of Conditional Probability

Let $(\Omega, \Pr)$ be a discrete probability space, and let $A \cap B \subseteq \Omega$ be events with $\Pr(B) \neq 0$. The conditional probability of $A$ given $B$, denoted $\Pr(A|B)$, is defined to be

$$
\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}
$$

← chopping down to $B$

← zooming in (scaling)
Example 1

We roll two fair dice. What is the probability the dice sum to 4, given that the first die is 3?

\[ \mathcal{S} = \{ (1,1), (1,2), \ldots, (6,6) \} , \quad |\mathcal{S}| = 36 \]

\[ A = \text{"sum to 4"} = \{ (1,3), (2,2), (3,1) \} \]

\[ B = \text{"first die = 3"} = \{ (3,1), (3,2), (3,3), (3,4), (3,5), (3,6) \} \]

\[ A \cap B = \{ (3,1) \} . \]

\[ \Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{1/36}{6/36} = \frac{1}{6} \]

What is the probability the first die is 3, given the sum is 4?

\[ \Pr(B \mid A) = \frac{\Pr(A \cap B)}{\Pr(A)} = \frac{1/36}{3/36} = \frac{1}{3} \]
Example 2

We flip two coins.

(a) What is the probability of two heads, given at least one head is shown?

(b) What is the probability of two heads, given the first coin showed heads?
Example 2

We flip two coins.

(a) What is the probability of two heads, given at least one head is shown?

(b) What is the probability of two heads, given the first coin showed heads?

For a: \( \mathcal{S} = \{HH, HT, TH, TT\} \). \( A = \{ \text{two heads} \} = \{HH\} \).

\( B = \{ \text{at least one head} \} = \{HH, HT, TH\} \)

\[
\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(HH)}{\Pr(B)} = \frac{1/4}{3/4} = \frac{1}{3}.
\]
Example 2

We flip two coins.

(a) What is the probability of two heads, given at least one head is shown?

(b) What is the probability of two heads, given the first coin showed heads?

For a: $S = \{HH, HT, TH, TT\}$. $A = \{\text{two heads}\} = \{HH\}$.

$B = \{\text{at least one head}\} = \{HH, HT, TH\}$

$$\Pr(A | B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(HH)}{\frac{3}{4}} = \frac{1}{3}.$$  

For b: $C = \{\text{first coin heads}\} = \{HT, HH\}$.

$$\Pr(A | C) = \frac{\Pr(A \cap C)}{\Pr(C)} = \frac{\Pr(HH)}{\frac{1}{2}} = \frac{1}{2}.$$
Exercises

1. Show $\Pr(A) = 0 \Rightarrow \Pr(A \mid B) = 0$.

2. Show $\Pr(A \cap B) > \Pr(A \mid B)$.

3. Find the probability a 5-card hand is all Spades, given that it is all the same suit.
Independence

Events
Intuition

"A is independent of B" should mean that whether or not A occurs has no bearing on whether or not B occurs.
Definition of Independence

Def Events $A, B$ are independent if

$$\Pr(A \cap B) = \Pr(A) \Pr(B).$$

If $\Pr(B) \neq 0$, then this implies

$$\Pr(A | B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(A) \Pr(B)}{\Pr(B)} = \Pr(A)$$

i.e. conditioning on $B$ does not change $\Pr(A)$.
Example 1

If we deal one random card atel
let \( A = \) "card is an Ace" and \( B = \) "card
is a spade" then

\[
\Pr(A) = \frac{1}{13}, \quad \Pr(B) = \frac{1}{4}, \quad \Pr(A \cap B) = \frac{1}{52}
\]

and \( \Pr(A \cap B) = \frac{1}{52} = \frac{1}{13} \cdot \frac{1}{4} = \Pr(A) \cdot \Pr(B). \)
Example 2

Roll two dice and let $A =$ "1st die is 3", $B =$ "sum is 7". Show $A, B$ are independent.
Example 2

Roll two dice and let $A = \text{"1st die is 3"}$,
$B = \text{"sum is 7"}$. Show $A, B$ are independent.

$\Pr(A) = \frac{1}{6}$ (done earlier)

$\Pr(B) = \Pr(\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\})$

$= \frac{6}{36} = \frac{1}{6}$

$\Pr(A \cap B) = \Pr(\{(3,4)\}) = \frac{1}{36} = \Pr(A) \cdot \Pr(B)$
Example 3

We often declare events to be independent as a modeling decision.

Suppose a coin is biased to show Heads with probability $\frac{3}{4}$. Then we'd model the probability of double Heads as $\frac{3}{4} \cdot \frac{3}{4} = \frac{9}{16}$ "because the tosses are independent".