UNIVERSITY of WASHINGTON

Conditional Probabilities

MATH/STAT 394: Probability I
Summer 2021 A Term

Introduction to Probability
D. Anderson, T.Seppäläinen, B. Valkó

§ 2.1

Aaron Osgood-Zimmerman

Department of Statistics

Quiz solution

Practice

Flip a fair coin 5 times. For each heads you pay me 1$ and for each tails I pay you 1$. Denote by X my net wining.

- What are the possible values for X?
- What is $\mathbb{P}(X = 3)$?

Solution

- Denote by H the number of heads in the 5 flips.
- The possible values of H are 0, 1, 2, 3, 4, 5
- The random variable of interest is $X = H - T = H - (5 - H) = 2H - 5$ so the possible values of X are -5, -3, -1, 1, 3, 5.
- Then we have that

$$\mathbb{P}(X = 3) = \mathbb{P}(H = 4) = \binom{5}{4} 2^{-4} 2^{-1} = 5 \cdot 2^{-5} \approx 0.16$$

  where $\binom{5}{4}$ is the number of to select 4 positions among the 5 possible for the heads and $2^{-4}2^{-1}$ is the probability of a tuple composed of 4 heads and 1 tail.

Quiz solution

Practice

Flip a fair coin 5 times. For each heads you pay me 1$ and for each tails I pay you 1$. Denote by X my net wining.

- What are the possible values for X?
- What is $\mathbb{P}(X = 3)$?

Solution

- Denote by H the number of heads in the 5 flips.
- The possible values of H are 0, 1, 2, 3, 4, 5
- The random variable of interest is $X = H - T = H - (5 - H) = 2H - 5$ so the possible values of X are -5, -3, -1, 1, 3, 5.
- Alternative computation: since the outcomes are equaly likely, and the number of ways to get 4 heads in 5 flips is $\binom{5}{4}$ we have

$$\mathbb{P}(H = 4) = \binom{5}{4} / \#\Omega = 5/32$$

Recap

Prob. definition

Probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Three axioms of probability $\mathbb{P}$

1. $0 \leq \mathbb{P}(A) \leq 1$, $A \in \mathcal{F}$
2. $\mathbb{P}(\Omega) = 1$
3. $\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$, if $A_i \cap A_j = \emptyset$ for $i \neq j$

Prob. calculus tools

- **Finite additivity:**
  - $\mathbb{P}(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mathbb{P}(A_i)$, if $A_i \cap A_j = \emptyset$ for $i \neq j$
- **Prob. of complement**
  - $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
  - $\mathbb{P}(B) = \mathbb{P}(B \cap A^c) + \mathbb{P}(B \cap A)$
- **Inclusion-exclusion formulas**
  - $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$
  - $$\begin{aligned} \mathbb{P}(A \cup B \cup C) = {} & \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) \\ & - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) \\ & + \mathbb{P}(A \cap B \cap C) \end{aligned}$$
  - See lecture 7 for general formula with $A_1, \ldots, A_n$

Outline

Conditional Probabilities

Motivation

Motivation

- The probability of getting a one when rolling a fair 6-sided die is 1/6
- Suppose you were given the extra information that the die roll was an odd number (hence 1, 3 or 5)
- Conditioned on this new information, intuitively, the probability of a one is now 1/3.
- Let's formalize this intuition

Conditional probability

Definition

Given two events A and B with $\mathbb{P}(B) > 0$, the ***conditional probability*** *of A given B is defined as*

$$\boxed{\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}}$$

- Make sure that $\mathbb{P}(B) > 0$ (o.w. you condition on something impossible)
- For equally likely outcomes,

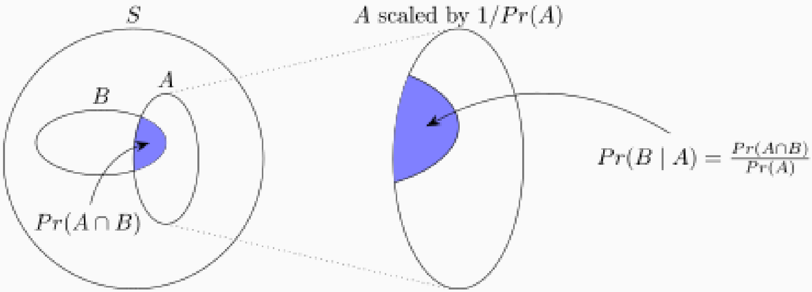
$$\boxed{\mathbb{P}(A \mid B) = \frac{\#(A \cap B)}{\#B}}$$

Note: Event $A \cap B$ is sometimes written AB.

Conditional probability intuition

Given two events A and B with $\mathbb{P}(A) > 0$, the **conditional probability** of B given A is defined as

$$\boxed{\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}}$$



Conditional probability

Example
Roll a die. $B = \{1 \text{ or } 3 \text{ or } 5\}$, $A = \{1\}$.

Prob. of A conditioned on B?

Solution

$$\mathbb{P}(\text{get } 1 \,|\, \text{roll is odd}) = \mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\text{get } 1)}{\mathbb{P}(\text{get } 1 \text{ or } 3 \text{ or } 5)} = \frac{1/6}{3/6} = \frac{1}{3}.$$

Conditional Probability

Exercise

Roll two fair 6-sided dice. Find the probability of the sum of two rolls is 7, conditioned on the first roll being 4.

Solution

- Denote two events $A = \{\text{Sum is } 7\}$ and $B = \{\text{First roll is } 4\}$.

$$\Omega = \{(i,j) : i = 1,\ldots,6, \quad j = 1,\ldots,6\}, \quad \#\Omega = 36.$$

$$AB = A \cap B = \{(4,3)\}, \quad B = \{(4,1),\ldots,(4,6)\}.$$

- Hence,

$$\mathbb{P}(A \cap B) = \#(AB)/\#\Omega = 1/36, \quad \mathbb{P}(B) = \#(B)/\#\Omega = 1/6,$$

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{1/36}{1/6} = \frac{1}{6}.$$

- Alternatively, since outcomes are equally likely,

$$\mathbb{P}(A \mid B) = \frac{\#(A \cap B)}{\#B} = \frac{1}{6} = \frac{1}{6}.$$

Some simple properties

Be careful

$$\mathbb{P}(A \mid B) \neq \mathbb{P}(B \mid A)$$

$\rightarrow \mathbb{P}(\text{raining} \mid \text{cloudy}) \neq \mathbb{P}(\text{cloudy} \mid \text{raining})$

Properties

1. $\mathbb{P}(A \mid A) = 1$.
2. $\mathbb{P}(A \mid \Omega) = \mathbb{P}(A)$.
3. $\mathbb{P}(A^c \mid A) = 0$.
4. $\mathbb{P}(A^c \mid B) = 1 - \mathbb{P}(A \mid B)$.

Generally:

Given B s.t. $\mathbb{P}(B) > 0$, then $\mathbb{P}(\cdot \mid B) : A \rightarrow \mathbb{P}(A \mid B)$ is a probability measure

Multiplication rule / Factorization

Proposition (Multiplication rule / Factorization[1])

- *By definition of conditional probability, the probability of $A \cap B$ is*

$$\boxed{\mathbb{P}(A \cap B) = \mathbb{P}(A \mid B)\mathbb{P}(B) = \mathbb{P}(B \mid A)\mathbb{P}(A)}$$

- *For three events A, B and C,*

$$\boxed{P(A \cap B \cap C) = \mathbb{P}(C \mid A, B)\mathbb{P}(A \cap B) = \mathbb{P}(C \mid A, B)\mathbb{P}(B \mid A)\mathbb{P}(A)},$$

where $\mathbb{P}(C \mid A, B)$ means $\mathbb{P}(C \mid A \cap B)$.

- *For n events $A_1, A_2, \ldots, A_n$:*

$$\boxed{\mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right) = \mathbb{P}(A_1)\mathbb{P}(A_2 \mid A_1)\mathbb{P}(A_3 \mid A_1, A_2)\cdots\mathbb{P}(A_n \mid A_1, \ldots, A_{n-1})}$$

[1] All definitions assume that the conditional probabilities are well-defined, i.e., that the conditioning event has a non-zero prob.

Multiplication rule / Factorization

Exercise

A box contains 8 red balls and 4 blue balls. Randomly draw 3 balls (X_1, X_2, X_3) without replacement. What is the probability of getting (R, B, B)?

Solution

1. *Direct calculation.*

$$\Omega = \{\text{3-tuples of } 1, \ldots, 12\}, \; A = \{X_1 = R, X_2 = B, X_3 = B\}.$$

By equally likely outcomes,

$$P(A) = \frac{\#A}{\#\Omega} = \frac{8 \cdot 4 \cdot 3}{(12)_3} = \frac{8 \cdot 4 \cdot 3}{12 \cdot 11 \cdot 10}.$$

2. *Multiplication rule.*

$$\begin{aligned} & P(X_1 = R, X_2 = B, X_3 = B) \\ = \; & P(X_1 = R) P(X_2 = B \mid X_1 = R) P(X_3 = B \mid X_1 = R, X_2 = B) \\ = \; & \frac{8}{12} \cdot \frac{4}{11} \cdot \frac{3}{10}. \end{aligned}$$

Law of total Probability

Exercise

Two urns. Urn I has 2 green balls and 1 red ball. Urn II has 2 red balls and 3 yellow balls.

Consider

1. Picking one of the two urns with equal prob.
2. Then sample one ball uniformly at random from the selected urn

What is the prob. of getting a red ball?

Solution

- Let $\{\text{urn I}\}$ denotes the event that we choose the urn I, denote $\{\text{urn II}\}$ same for urn II and denote $\{\text{red}\}$ the event that we select a red ball
- Decompose $\{\text{red}\}$ as

$$\begin{aligned}\mathbb{P}(\{\text{red}\}) &= \mathbb{P}(\{\text{red}\} \cap \{\text{urn I}\}) + \mathbb{P}(\{\text{red}\} \cap \{\text{urn II}\}) \\ &= \mathbb{P}(\text{red} \mid \text{urn I})\mathbb{P}(\text{urn I}) + \mathbb{P}(\text{red} \mid \text{urn II})\mathbb{P}(\text{urn II}) \\ &= \frac{1}{3} \cdot \frac{1}{2} + \frac{2}{5} \cdot \frac{1}{2} = \frac{11}{30}\end{aligned}$$

Law of total probability

Law of total probability (simple version)

General version of the reasoning used in the last example:[2]

$$\boxed{\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c) = \mathbb{P}(A \mid B)\mathbb{P}(B) + \mathbb{P}(A \mid B^c)\mathbb{P}(B^c)}$$

Notes:

- This is a recurring approach throughout probability.
- Just decompose a complex event into simpler disjoint pieces
- Here we used (B, B^c) to split A
- This can be generalized to partitions.

[2] Assuming that the cond. prob are well defined, i.e. $\mathbb{P}(B) > 0$ and $\mathbb{P}(B^c) > 0$.

Law of total probability

Definition

*Subsets/events $B_1, \ldots, B_n$ of Ω form a **partition** of Ω if*

1. *they are pairwise disjoint, i.e., $B_i \cap B_j = \emptyset$ for any $i \neq j$*
2. *they cover the sample space, i.e., $\cup_{i=1}^n B_i = \Omega$*

Proposition (Law of total probability)

Let $B_1, \ldots, B_n$ be a partition of Ω with $\mathbb{P}(B_i) > 0$ for all i.

For any event A,

$$\begin{aligned}\mathbb{P}(A) &= \sum_{i=1}^n \mathbb{P}(A \cap B_i) \quad \textit{(Finite additivity)} \\ &= \sum_{i=1}^n \mathbb{P}(A \mid B_i)\mathbb{P}(B_i) \quad \textit{(Multiplication rule)}\end{aligned}$$

Venn Diagrams Representations

The partitioning of A by B can be represented as follows

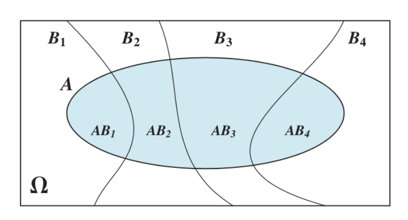


Figure from Introduction to probability, D. Anderson, T. Seppäläinen, B. Valkò

Quiz next lecture

Practice (Example from a Wall Street job interview)
Let us play a Russian roulette. You are tied to your chair. Here's a gun, a revolver.

Here's the barrel of the gun, six chambers, all empty. Now watch me as I put two bullets into the barrel, into two adjacent chambers. I close the barrel and spin it. I put a gun to your head and pull the trigger.

Click. Lucky you!

Now I'm going to pull the trigger *one more time*. Which would you prefer: that I spin the barrel first or that I just pull the trigger?