

Introduction

Book: Not mandatory

- Uploaded relevant exercises
- Use another book or Google

Prerequisites: Recap today

Expect you to fill in the blanks

Exam: 3 hour in Wiseflow

Documentation must be uploaded

Must be .ipynb format

Tools: Python 3

Jupyter Notebook

↳ VSCode

↳ Jupyter lab

↳ DataSpell (Jetbrains)

↳ Google Colab

itslearning: I use in list mode

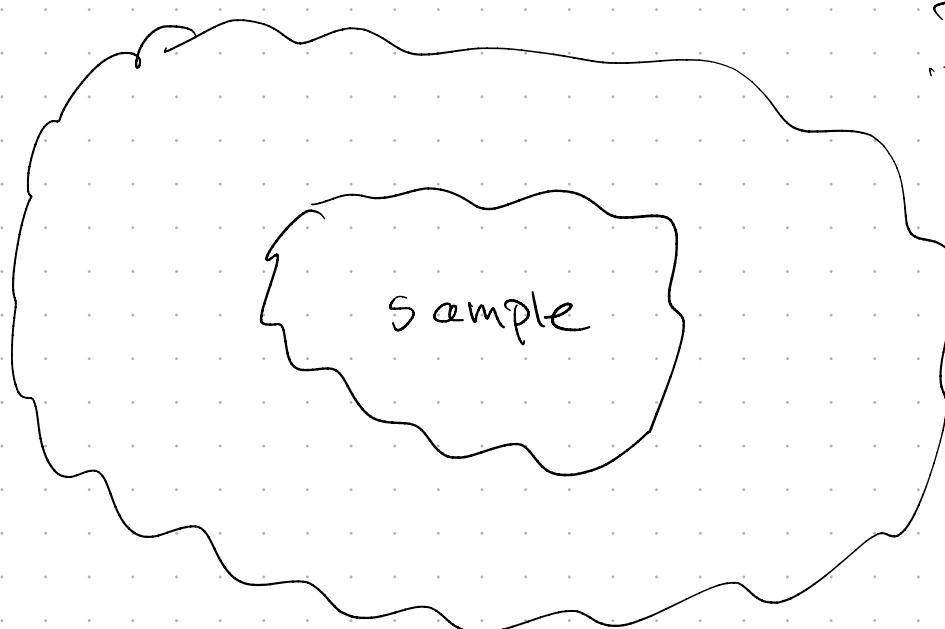
Session 0 will contain all material not associated with a specific session!

Wiseflow: You will receive multiple Wiseflow assignment

Why Probability and statistics?

- Stochastic = Random
- Natural Processes \rightarrow Variability \rightarrow Uncertainty
 - \hookrightarrow Statistics is the study of how to deal with uncertainty
 - \hookrightarrow Probability is the unit in which we measure uncertainty

Samples and Populations:



Population:
all Android
Devices

Sample = Representative subset of
Population

Sample Size \rightarrow Big Data: $E = \frac{Z \cdot \sigma}{\sqrt{n}}$

Scales and Measurements:

Nominal: Values have no quantitative significance: 0 = male, 1 = female
Is $1 > 0$?

↳ Groups, classes, categories
e.g. gender, color, jobs, etc.

Ordinal: Values are comparable, but difference is not known

↳ Order matters: Ranking.

Interval: Values are comparable and difference and distance matter

↳ Zero is assigned arbitrarily

- * Temperature
- * Time of day
- * Dates
- * Likert scale

Ratio: Same as interval but with a "natural" zero:

- * Height
- * Weight
- * Income
- * Number of children

Random Experiments

- Process where something uncertain is observed
- An outcome is the result of a random experiment
- A sample space S is the set of all possible outcomes.

Toss a coin:

$$S = \{H, T\}$$

Guess a bit string:

$$S = \{0, 1\}^n$$

Roll a die:

$$S = \{1, 2, \dots, 6\}$$

Price of item:

$$S = \{0, \dots, \infty\}$$

Goals scored in football match:

$$S = \{0, 1, \dots, \infty\}$$

Goals scored in Italian football:

$$S = \{0, 1\}$$

- An event A is a subset of the sample space.

Probability

Objective / classical

- Based on equally likely events
- long run relative frequency
- same for all observers

Empirical

- Based on observation
- Relative frequency of large amount of observations

Subjective

- Based on personal belief, experience, prejudice, etc.
- Different for all observers

Axioms of Probability

1. For any event A , $P(A) \geq 0$

2. $P(S) = 1$

3. If A_1, A_2, A_3, \dots are disjoint

$$P(A_1 \cup A_2 \cup A_3 \dots) = P(A_1) + P(A_2) + P(A_3) + \dots$$

Notation:

$$P(A \cap B) = P(A \text{ and } B) = P(A, B)$$

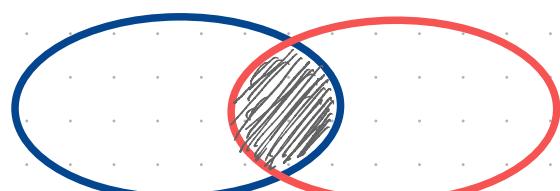
$$P(A \cup B) = P(A \text{ or } B)$$

$$P(\bar{A}) = P(A') = P(A^c) = P(\text{not } A)$$

Rules of Probability

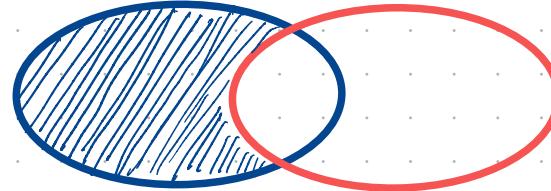
a) $P(A^c) = 1 - P(A)$

$A \cap B$



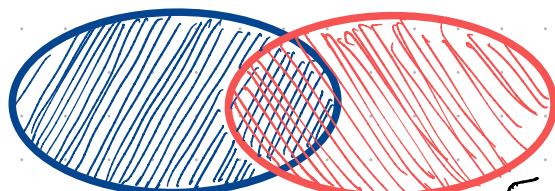
b) $P(\emptyset) = 0$

$A - B$



c) $P(A) \leq 1$

$A \cup B$



d) $P(A - B) = P(A) - P(A \cap B)$

e) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

f) $A \subset B \rightarrow P(A) \leq P(B)$

Is a subset of

Example:

1. 60% Chance of rain today: $P(A) = 0.6$
2. 50% Chance of rain tomorrow: $P(B) = 0.5$
3. 30% Chance of no rain either day: $P(A^c \cap B^c) = 0.3$

a. Probability of rain either day:

$$\begin{aligned} P(A \cup B) &= 1 - P(A \cap B)^c && \text{De Morgan's} \\ &= 1 - P(A^c \cap B^c) && \text{law} \\ &= 1 - 0.3 = \underline{\underline{0.7}} \end{aligned}$$

b. Rain both days:

$$\begin{aligned} P(A \cap B) &= P(A) + P(B) - P(A \cup B) \\ &= 0.6 + 0.5 - 0.7 \\ &= \underline{\underline{0.4}} \end{aligned}$$

c. Rain today, not tomorrow:

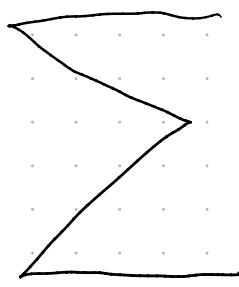
$$\begin{aligned} P(A - B) &= P(A) - P(A \cap B) \\ &= 0.6 - 0.4 \\ &= \underline{\underline{0.2}} \end{aligned}$$

d. Rain today or tomorrow, not both

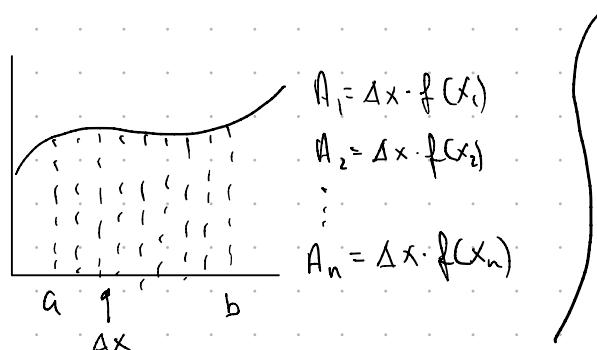
$$\begin{aligned} P((A \cup B) \setminus (A \cap B)) &= P(A \cup B) - P(A \cap B) \\ &= 0.7 - 0.4 = \underline{\underline{0.3}} \\ \text{OR} \\ &= P(A - B) + P(B - A) \\ &= 0.2 + (0.5 - 0.4) = \underline{\underline{0.3}} \end{aligned}$$

Discrete Probability Models

Discrete



Continuous



$$A \approx A_1 + A_2 + A_3 + \dots + A_n \approx \sum_a^b \Delta x \cdot f(x_i) = \int_a^b f(x) dx$$

$S \rightarrow \int \rightarrow \lim_{\Delta x \rightarrow 0}$

If a sample space is countable, we use a discrete probability model

If $A \subset S$:

$$P(A) = P\left(\bigcup_{S_i \in A} \{S_i\}\right) = \sum_{S_i \in A} P(S_i)$$

Example:

Assume you win $K-2$ Kroner with $P = \frac{1}{2^K}$, $K \in \mathbb{N}$

a. loose 1: $K-2 = -1 \Rightarrow K = 1$ so $P_1 = \frac{1}{2}$

b. loose 0: $K-2 = 0 \Rightarrow K = 2$ so $P_0 = \frac{1}{4}$

c. Win 1: $K-2 = 1 \Rightarrow K = 3$ so $P_1 = \frac{1}{8}$

d. Win 2: $K-2 = 2 \Rightarrow K = 4$ so $P_2 = \frac{1}{16}$

e. Win 3: $K-2 = 3 \Rightarrow K = 5$ so $P_3 = \frac{1}{32}$

f. Win [1; 3]: $P_1 + P_2 + P_3 = \frac{7}{32}$

g. More than 2: $P(3) + P(4) + P(5) + \dots$

$$= \frac{1}{16} + \frac{1}{32} + \frac{1}{64} + \dots$$

Conditional Probability

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) > 0$$

Some important consequences:

$$\rightarrow P(A^c|B) = 1 - P(A|B)$$

$$\rightarrow P(\emptyset|B) = 0$$

$$\rightarrow \text{If } B \subset A, \quad P(A|B) = 1 = \frac{P(A \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$$

$$\rightarrow \text{If } A \subset B, \quad P(A|B) = \frac{P(A)}{P(B)}$$

Chain Rule

$$P(A_1 \cap A_2 \cap A_3 \dots A_k) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1, A_2) \cdot \dots \cdot P(A_k|A_{k-1}, A_{k-2}, \dots)$$

Independence

$$P(A|B) = P(A) \text{ iff. } A \text{ and } B \text{ are independent}$$

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

}

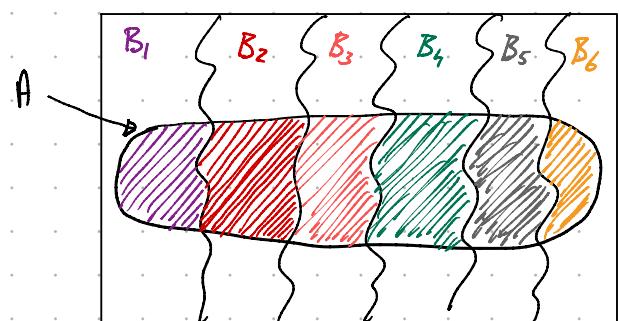
$$P(A \cap B) = P(A|B) \cdot P(B) = P(A) \cdot P(B)$$

Law of total probability

$$P(A) = \sum P(A \cap B_i)$$

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + \dots$$

$$P(A) = \sum P(A|B_i) \cdot P(B_i)$$



Bages' Rule

$$P(A|B) \cdot P(B) = P(A \cap B) = P(B|A) \cdot P(A)$$

$$P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)}$$

$$P(B_j|A) = \frac{P(A|B_j) \cdot P(B_j)}{\sum_i^k P(A|B_i) \cdot P(B_i)}$$

$$P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A|B) \cdot P(B) + P(A|\bar{B}) \cdot P(\bar{B})}$$

Conditional Independence (not recap)

Two events A and B are said to be conditionally independent given an event C if

$$P(A \cap B|C) = P(A|C) \cdot P(B|C), P(C) > 0$$

Understanding Conditional independence:

$P(A|B,C)$ = A given B and C

If A and B are conditionally independent,

then $P(A|B,C)$ does not depend on B and we get $P(A|B,C) = P(A|C)$

and also $P(A,B|C) = P(B|C)$, i.e.

A and B are independent given C.

Example

A box contains two coins: a regular coin and one fake two-headed coin ($P(H) = 1$). I choose a coin at random and toss it twice. Define the following events.

- A= First coin toss results in an H .
- B= Second coin toss results in an H .
- C= Coin 1 (regular) has been selected.

Find $P(A|C)$, $P(B|C)$, $P(A \cap B|C)$, $P(A)$, $P(B)$, and $P(A \cap B)$. Note that A and B are NOT independent, but they are *conditionally* independent given C .

Regular Coin: $P(H) = \frac{1}{2}$, $P(T) = \frac{1}{2}$

Fake Coin : $P(H) = 1$, $P(T) = 0$

$$P(A|C) = \frac{1}{2}$$

$$P(B|C) = \frac{1}{2}$$

$$P(A \cap B|C) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

$$\begin{aligned} P(A) &= P(A|C) \cdot P(C) + P(A|\bar{C}) \cdot P(\bar{C}) \\ &= \frac{1}{2} \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} \\ &= \frac{3}{4} \end{aligned}$$

$$\begin{aligned} P(B) &= \text{Same as } P(A) \\ &= \frac{3}{4} \end{aligned}$$

$$\begin{aligned} P(A \cap B) &= P(A \cap B|C) \cdot P(C) + P(A \cap B|\bar{C}) \cdot P(\bar{C}) \\ &= P(A|C) \cdot P(B|C) \cdot P(C) + P(A|\bar{C}) \cdot P(B|\bar{C}) \cdot P(\bar{C}) \\ &= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} + 1 \cdot 1 \cdot \frac{1}{2} \\ &= \frac{1}{8} + \frac{1}{2} = \underline{\underline{\frac{5}{8}}} \end{aligned}$$

Random Variables

A random variable X is a function from the sample space to \mathbb{R} :

$$X: S \rightarrow \mathbb{R}$$

The range R_X is the set of all possible values of X

Ex:

- Toss a coin 100 times. Let X be the number of heads: $R_X = \{0, 1, \dots, 100\}$
- Toss a coin until the first head. Let Y be the number of tosses until first head:
$$R_Y = \mathbb{N}_+$$
- Let Z denote the number of times you have to stop at a red light on your way to school:

$$P(Z=0)$$

$$P(Z=1)$$

$$P(Z \leq 2)$$

:

etc.