

# 1 Lecture 1

## 1.1 Probability vs. Statistics

- Statistics is probability formulate differently
- Statistics applies to probability to extract information about data and recommend what decisions should be made

**Probabilistic:** You have a fair coin, you flip it 100 times, what is the probability to getting more than a 70% heads (We expect a small number).

**Statistical:** You have a coin of unknown fairness, you flip it 100 times and get 70% heads. Your job would be to draw conclusions about that coin.

The main difference is that you know the conditions when you are deriving probability of a particular outcome. In statistics, you know the outcome and you are trying to obtain conclusions about the underlying relations, models, and hidden patterns in the data.

## 1.2 Frequency vs. Bayes

Consider a simple coin toss and assume you toss it  $n$  times. Let  $nh$  be the number of heads and  $nt = n - nh$  be the number of tails.

The **Frequentist** estimates the probability of heads as  $Ph = nh/n$ . That is the frequency of the "head" event.

The **Bayesian** attempts to find prior knowledge such as where was the coin minted and built. This "prior" information is then used for analyzing the data.

## 1.3 Counting and Sets

We need to determine all the possible all the outcomes in order to determine probabilities. To accomplish this we would need to understand sets, unions, intersections, and complements. Also aspects of combinatorics calculus is needed in order to analyze problems eg. selecting  $n$  out of  $m$  objects.

Ex. Toss a fair coin 4 times. What is the probability to get exactly two tails?

We can approach this using the counting approach by finding all the possible outcomes.

$S = \{ \text{HHHH, HHHT, HHTH, HHTT, HTHH, HTHT, HTTH, HTTT, THHH, THHT, THTH, THTT, TTHH, TTHT, TTTH, TTTT} \}$

Then we find the set of favorable outcomes

$A = \{ \text{HHTT, HTHT, HTTH, THHT, THTH, TTHH} \}$

So the probability of the event to exactly two tails in 4 tosses is the number of favourable outcomes divided by the total outcomes =  $6/16$ .

But this type of approach becomes impossible once the space grows. Ex. if we had 100 tosses.

## 1.4 Formalizing Sets/Counting

A **set** is a collection of elements: Ex. patients with blood pressure over 140, people with brown hair, etc.

We use  $x \in S$  to denote **element**  $x$  belonging to the set  $S$ : Ex. John has brown hair, hence John is an element in the set of people with brown hair.

We say a set  $A$  is a **subset** of  $S$  if all the elements in  $A$  belongs to  $S$ : Ex. we can pick 6 people from the people with brown hair set and make a new set. This new set would be a subset of the original set of people with brown hair.

An **empty set** is the set with no elements denoted as  $\emptyset$

## 1.5 Set Operations

**Complement:** The complement of  $A$  in  $S$  is all the elements in  $S$  that are not in  $A$ . We denote this as  $S - A$  or  $A^c$ .

Ex. Let  $B$  be the set of all the people with brown hair and shorter than 6 feet. The complement of this would be all the people who do not have brown hair or are taller than 6 feet.

**Union:** The union of  $A$  and  $B$  is the set of all the elements that are in  $A$  or  $B$  (or in both). We denote this as:  $A \cup B$ .

**Intersection:** The intersection of  $A$  and  $B$  is the set of all elements that are in both  $A$  and  $B$ . We denote this as:  $A \cap B$ .

**Disjoint:** Two sets are called disjoint if they have no mutual elements

**Difference:** The difference of  $A$  and  $B$  is the set of all elements in  $A$ , but are not in  $B$  and is denoted as  $A \setminus B$ .

Ex. Consider a set of 10 cities {Berlin, Munich, Madrid, Barcelona, Beijing, Shanghai, Portland, Washington DC, Ottawa, Toronto}

Consider two subsets: 1)  $M$  is a set of capitols and 2)  $N$  is a set of cities in Europe

$M = \{\text{Berlin, Madrid, Beijing, Washington DC, Ottawa}\}$

$N = \{\text{Berlin, Munich, Madrid, Barcelona}\}$

- Intersection: {Berlin, Madrid}
- Union: {Berlin, Munich, Madrid, Barcelona, Beijing, Washington DC, Ottawa}

- Complement of M in S : {Munich, Barcelona, Shanghai, Portland, Toronto}
- Difference = {Beijing, Washington DC, Ottawa}

## 1.6 De Morgans Laws

$$(A \cup B)^c = A^c \cap B^c \quad (1)$$

$$(A \cap B)^c = A^c \cup B^c \quad (2)$$

## 1.7 Proof of De Morgans Laws

$$\begin{aligned} x &\in (A \cap B)^c \\ \iff x &\notin (A \cap B) \\ \iff x &\in A^c \cup x \in B^c \end{aligned}$$

$$\begin{aligned} x &\in (A \cup B)^c \\ \iff x &\notin (A \cup B) \\ \iff x &\in A^c \cap x \in B^c \end{aligned}$$

## 1.8 Venn Diagrams



Figure 1: Set  $S$  of all possible outcomes

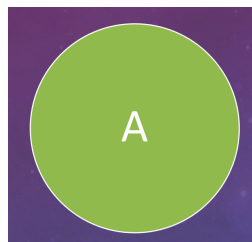


Figure 2: The set of all the outcomes in  $A$

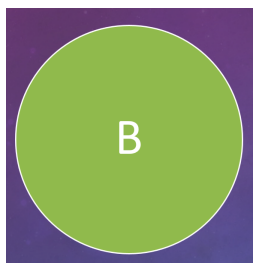


Figure 3: The set of all the outcomes in  $B$



Figure 4: Union of set  $A$  and set  $B$



Figure 5: Intersection of  $A$  and  $B$



Figure 6: Complement of  $A$  or  $A^c$

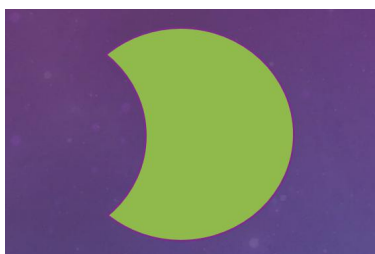


Figure 7:  $B - A$  or  $B \setminus A$

## 1.9 Set Product and Additional Set Terminology

$$S \times T = \{(s, t) \in S \text{ and } t \in T\}$$

Ex.  $\{1, 2\} \times \{3, 4, 5\} = \{(1, 3), (1, 4), (1, 5), (2, 3), (2, 4), (2, 5)\}$

Cardinal number is the number of elements in the set. Cardinal number is denoted as  $\#s$  or  $card\{S\}$  or  $|S|$ .

Inclusion - Exclusion principle:  $|A \cup B| = |A| + |B| - |A \cap B|$

We subtract  $|A \cap B|$  to ensure we do not double count.