

Class Notes

STAT410

Wenyu Chen

Summer 2022

1 | Axioms of Probabilities

Definition 0.0.0.0.1

Experiment: is repeatable task with well defined outcomes.

Sample Space: is a collection of all possible outcomes of the experiment.

Event: is a subset of the sample space.

Example: Suppose we toss a coin three times, assume coins lands on H or T.

Experiment: tossing the coin three times, noting the outcomes.

Sample Space: {HHH, HHT, HTH, THH, HTT, THT, TTH, TTT }

Example of events: $E_1 = \text{getting all heads} = \{HHH\} \subseteq S$

$E_2 = \text{getting exactly one H and one T} = \{\} \subseteq S$

$E_3 = \text{getting at least two heads} = \{HHH, HHT, HTH, THH\} \subseteq S$

We want to assign a number to each event, which is a measure of the chance or probabilities that this event happened. Our goal is to understand the process of assigning probabilities to events.

Example: Die Roll

Experiment: Roll a six sided dice two times

Sample Space: $\{(1, 1), (1, 2), \dots, (1, 6), (2, 1), \dots, (2, 6), \dots, (6, 6)\}$

Example of events:

$E_1 = \text{at least one six} = \{(1, 6), (2, 6), (3, 6), (4, 6), (5, 6), (6, 6), (6, 1), (6, 2), (6, 3), (6, 4), (6, 5)\}$

$|E_2| = 10$

$E_2 = \text{same numbers on both rolls} = \{(1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6)\}$

$E_4 = \{(1, 5), (3, 4)\}$

Note:

Simple Event: is the event with exactly one outcomes. It is the cardinality of the sample space.

Number of events: Assume the sample space is finite of size n, the number of events is the cardinality of all possible outcomes of $S = 2^n$

Set in context

Let an event E be describes as a subset of the sample space S.

Let E, F two events be given.

$E \cap F$ is a new event that corresponds to outcomes in both E and F

$E \cup F$ is a outcome a new event include outcomes in E or F

E^c is a new event that include outcomes where E doesn't happen.

When we say an event E has happened, we means that the outcome ω of the experiment lies inside E

Example: Tossed a coin three times. In one run of the experiment, the result is HHT.

$HHT \in E_1 = \text{at least two heads}$, we said E_1 has happened.

Definition 0.0.0.0.2

A sigma algebra \mathcal{B} for a set s is a collection of subset of S that satisfies:

1. $\emptyset \in \mathcal{B}$
2. If $A \in \mathcal{B}$ then $A^c \in \mathcal{B}$ (closed under complement)
3. If $\{A_i\}_{i=1}^{\infty} \in \mathcal{B}$ then $\cap_{i=1}^{\infty} A_i \in \mathcal{B}$ (closed under)

Sigma algebra is a collection of events to which we want to assign the probabilities.

Axioms for a probability function

Suppose we are given the pair (S, \mathcal{B}) which S represents the sample space and \mathcal{B} is a sigma algebra \mathcal{B} .

Definition 1.0.0.0.1

A probability function P satisfies the following:

$$\begin{aligned} P : \mathcal{B} &\rightarrow \mathbb{R} \\ E &\mapsto P(E) \end{aligned}$$

1. $P(A) \geq 0$ for all $A \in \mathcal{B}$
2. $P(S) = 1$
3. If A_1, A_2, A_3, \dots are mutually disjoint sets in \mathcal{B} then $P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$

Theorem 1.1

Suppose the sample space S is countable. To define a probability function on $(S, \mathcal{B} = \mathcal{P}(S))$, we do the following

1. Find a seq $\{p_i\}_{i=1}^{\infty}$ such that (i) $0 \leq p_i \leq 1$ and (ii) $\sum_{i=1}^{\infty} p_i = 1$
2. Define $P(\{s_i\}) = p_i$
3. For any $E \in \mathcal{B}, E = \{s_{i1}, s_{i2}, \dots, s_{ij}\}$

Example: Probability Functions Examples

Coin Tosseo:

Experiment: tossing the coin three times, noting the outcomes.

Sample Space: $\{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$

To get a probability function, we will need to work with a sigma algebra. Suppose $\mathcal{B}_1 = \mathcal{P}(S)$, note that the cardinality of \mathcal{B}_1 is 256.

There are infinitely many ways to choose the sequence p_i . One way is to choose $p_i = \frac{1}{8}$ for all i . Another way is to set $p_1 = 1$ and rest to 0.

Tree Diagram is a graph that describes the flow of the outcomes of each steps in an experiment.

Definition 1.1.0.0.1

Conditional Probability is $P(A | B) = P(A \text{ happens given that } B \text{ has already happened})$

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

And by **Multiplication Principle**,

$$\begin{aligned} P(A \cap B) &= P(A | B) \times P(B) \\ &= P(B | A) \times P(A) \end{aligned}$$

Assumption: 1) Probability of events when all outcomes in the sample space are equally likely. 2) Sample space is finite.

Proposition 1.1.1

If every outcome in the sample space is equally likely, we can calculate the probability of $E \subseteq S$ as follow:

$$P(E) = \frac{n(E)}{n(S)}$$

where $n(E)$ is the number of outcomes in E .

Example: Suppose we toss a coin that $p(H) \in (0, 1)$ two times. The sample space is $S = \{HH, HT, TH, TT\}$. If the coin is fair, then $P(HH) = P(HT) = P(TH) = P(TT) = 0.25$

Theorem 1.2

Fundmental Theorem of Counting: Suppose a task T can be performed as a sequence of subtasks: $T_1, T_2, T_3, \dots, T_k$. And each $n_1, \dots, n_2, n_3, \dots, n_k$ is number of ways to perform T_i Then the total number ways to perform the task T is

$$n_1 \times n_2 \times n_3 \times \dots \times n_k$$

Typically we will have to select k objects from n distinct objects.

Example: $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$, we might be interested in knowing the total number of ways one can choose 4 digits from this 10 digits.

	Without Replacement	With Replacement
Order Matters	(1,2,4,5) different from (1,5,2,4) (1,1,2,5) is not possible	(1,2,4,5) different from (1,5,2,4) (1,1,2,5) is possible
Order Does Not Matters	(1,2,3,4) is same as (4,3,2,1) (1,1,2,5) not possible	(1,2,3,4) is same as (4,3,2,1) (1,1,2,4) is possible

1. Without replacement and order matters

Use the fundamental theorem of counting, we divide T , which is select k digits from a set of n distinct objects divide into

$$T : T_1 \rightarrow T_2 \rightarrow T_3 \rightarrow \dots \rightarrow T_k$$

where T_i is select i th object. Then, we will got

$$n \times (n - 1) \times (n - 2) \times (n - 3) \times \dots \times (n - k + 1)$$

Then, we got

$${}_nP_k = \frac{n!}{(n - k)!}$$

2. Without replacement and order does not matter

T = choose k objects from n distinct objects where order does not matter and without replacement.

$$T : T_1 \rightarrow T_2$$

T_1 is choose k objects where order matters and without replacement

T_2 is to get rid of all the times we have double counted.

ways to do $T_1 = {}_nP_k = \frac{n!}{(n-k)!}$

ways to do T_2 = number of arrangements of k objects = $k!$

Then we got

$${}_nC_k = \binom{n}{k} = \frac{n!}{(n - k)!k!}$$

3. With replacement and order does not matter

T = Choose k objects where order does not matter and with replacement.

We keep track of how many times a given object repeats in the selection and the total number of objects in the selection is equal to k .

Which is the same as choosing $n - 1$ walls in a set of $(n + k - 1)$, we got

$${}^{n+k-1}C_k = {}^{n+k-1}C_{n-1}$$

Example: $\{1, 2, 3, 4\}$, $k = 10$, We are selecting 10 objects from $\{1, 2, 3, 4\}$ with replacement and order does not matter.

We only care about how many times each number shows up since order does not matter and the total objects in selection are 10. To achieve that, we setup 13 spots. Such that, the extra position is how many times the given number repeats in the selection.

Such that, the

2 | Conditional Probability

Recall that given two events A and B, the conditional probability is

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

Definition 0.0.0.0.1

WE say two events A and B are independent if and only if

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

$$P(A | B) = P(A)$$

$$P(B | A) = P(B)$$

This means that B has happened will not affect the probability of A happening. Equivalent to saying that $P(A | B) = P(A)$

Note: $P(A \cap B)$ is the probability that A and B happen simultaneously.

Theorem 0.1

$$\begin{aligned} P(A | B) &= \frac{P(A \cap B)}{P(B)} \\ &= \frac{P(B | A) \cdot P(A)}{P(B)} \end{aligned}$$

and

$$\begin{aligned} P(B | A) &= \frac{P(B \cap A)}{P(A)} \\ &= \frac{P(A | B) \cdot P(B)}{P(A)} \end{aligned}$$

Exercise: Now if A and B are independent, we want to check if A and B^c is independent, which is checking if $P(A \cap B^c) = P(A)P(B^c)$.

Example:

Experiment:

Step1: Toss a coin with $P(H) = 0.4$

Step2: If we got head, we will roll a fair dice.

Else, we will roll a dice with $P(1) = P(2) = \dots = P(5) = 0.1$ and $P(6) = 0.5$

Question: Given that they observed a 6, what is the probability that a fair dice was rolled?

So we need to calculate $P(H | 6)$ which is the probability of rolling a fair dice (toss 6) given that 6 was observed.

f