CS 107, Probability, Spring 2020
Lecture 18

Michael Poghosyan
mpoghosyan@aua.am

AUA

28 February 2020

Content

- Review Session

Experiment, Outcomes and the Sample Space

Everything starts from:

- An Experiment (Random Experiment)
- An Outcome in an Experiment
- The Sample Space of the Experiment
- An Event in an Experiment

Probability (Measure) Definition

Assume we have an Experiment, Ω is its Sample Space and $\mathcal{F}$ is the set of all Events.

Probability Measure Definition

A function $\mathbb{P} : \mathcal{F} \to \mathbb{R}$ is called a **Probability Measure** on $(\Omega, \mathcal{F})$, if it satisfies the following axioms:

P1. For any $A \in \mathcal{F}$, $\mathbb{P}(A) \geq 0$;

P2. $\mathbb{P}(\Omega) = 1$;

P3. For any sequence of pairwise mutually exclusive (disjoint) events $A_n \in \mathcal{F}$, i.e., for any sequence $A_n \in \mathcal{F}$ with $A_i \cap A_j = \varnothing$ for $i \neq j$, we have

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n).$$

Properties of the Probability Measure

1. $\mathbb{P}(\varnothing) = 0$;
2. if $A, B \in \mathcal{F}$ are mutually exclusive events, i.e., if $A \cap B = \varnothing$, then
$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B);$$
3. for any event $A \in \mathcal{F}$,
$$\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A);$$
Here $\overline{A} = A^c = \Omega \setminus A$.

Properties of the Probability Measure

4. If $A_1, A_2, ..., A_n \in \mathcal{F}$ are pairwise disjoint (mutually exclusive), i.e., if $A_i \cap A_j = \varnothing$ for $i \neq j$, then

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i);$$

5. for any events $A, B \in \mathcal{F}$ (not necessarily disjoint),

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B);$$

Properties of the Probability Measure

6. If $A_1, A_2, ..., A_n \in \mathcal{F}$ are events, not necessarily disjoint, then

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + ... + \mathbb{P}(A_n) -$$

$$-\mathbb{P}(A_1 \cap A_2) - \mathbb{P}(A_1 \cap A_3) - ... - \mathbb{P}(A_{n-1} \cap A_n) +$$

$$+\mathbb{P}(A_1 \cap A_2 \cap A_3) + \mathbb{P}(A_1 \cap A_2 \cap A_4) + ... + \mathbb{P}(A_{n-2} \cap A_{n-1} \cap A_n) - ...$$

$$... + (-1)^{n-1}\mathbb{P}(A_1 \cap A_2 \cap ... \cap A_n).$$

This is the general version of the previous property, and is called the inclusion-exclusion principle.

Properties of the Probability Measure

7. for any events $A, B \in \mathcal{F}$

$$\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A \cap B);$$

8. if $A, B \in \mathcal{F}$ are two events such that $A \subset B$, then

$$\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A);$$

9. if A and B are two events such that $A \subset B$, then

$$\mathbb{P}(A) \leq \mathbb{P}(B);$$

10. for any event $A \in \mathcal{F}$,

$$0 \leq \mathbb{P}(A) \leq 1;$$

Classical Probability Models: Finite Sample Spaces

The Finite Sample Space Probability Model is of the form:

Outcome	ω_1	ω_2	...	ω_n
$\mathbb{P}(\{\omega_k\})$	p_1	p_2	...	p_n

with $p_1 + p_2 + ... + p_n = 1$, $p_k \geq 0$, $k = 1, ..., n$, and we also define

$$\mathbb{P}(A) = \sum_{\omega_i \in A} p_i,$$

and add $\mathbb{P}(\varnothing) = 0$.

Classical Probability Models: Countably Infinite Sample Spaces

This model looks like this:

Outcome	ω_1	ω_2	ω_3	...	ω_n	...
Probability	p_1	p_2	p_3	...	p_n	...

where $p_k \geq 0$ for any k, and $\sum_{k=1}^{\infty} p_k = 1$.
Now, we define for any nonempty event $A \in \mathcal{F}$ (i.e., for any nonempty subset $A \subset \Omega$),

$$\mathbb{P}(A) = \sum_{\omega_k \in A} p_k,$$

and also $\mathbb{P}(\varnothing) = 0$.

Equally Likely Outcomes

Now assume we have a Discrete Model with finitely many outcomes:

$$\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$$

and assume

$$\mathbb{P}(\{\omega_1\}) = \mathbb{P}(\{\omega_2\}) = ... = \mathbb{P}(\{\omega_n\}) = \frac{1}{n}.$$

In the table form:

Outcome	ω_1	ω_2	ω_3	...	ω_n
Probability	$\frac{1}{n}$	$\frac{1}{n}$	$\frac{1}{n}$	...	$\frac{1}{n}$

In this case,

$$\mathbb{P}(A) = \frac{\text{number of elements favorable for the event } A}{\text{total number of possible outcomes}} = \frac{\#A}{\#\Omega}.$$

Classical Prob Models: Geometric Probabilities

- Our Experiment's Sample Space is $\Omega \subset \mathbb{R}^n$;
- If $A \subset \Omega$ is an Event, then we define

$$\mathbb{P}(A) = \frac{measure(A)}{measure(\Omega)}.$$

Here *measure* $=$

- $\#$ (number of elements) in 0D;
- length in 1D
- area in 2D
- volume in 3D
- volume in 4D
- ....

Conditional Probabilities

Assume Ω is our Experiment's Sample Space, and A, B are two events.

Conditional Probability

The conditional probability of A given B (or the probability of A under the condition of B) is defined to be

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

Interpretation:

Properties of Conditional Probabilities

Assume $B \subset \Omega$ is a fixed event with $\mathbb{P}(B) > 0$. Then

a. For any event A,
$$\mathbb{P}(A|B) \geq 0;$$

b. $\mathbb{P}(B|B) = 1$;

c. If A is an event, then
$$\mathbb{P}(\overline{A}|B) = 1 - \mathbb{P}(A|B);$$

d. If $A_1, ..., A_n$ are some **mutually disjoint** events, then
$$\mathbb{P}(A_1 \cup A_2 \cup ... \cup A_n|B) = \mathbb{P}(A_1|B) + \mathbb{P}(A_2|B) + ... + \mathbb{P}(A_n|B);$$

e. If $A_1, ..., A_n, ...$ are some **mutually disjoint** events, then
$$\mathbb{P}(\cup_{n=1}^{\infty} A_n|B) = \sum_{n=1}^{\infty} \mathbb{P}(A_n|B);$$

Properties of Conditional Probabilities, Cont'd

f. If A_1, A_2, B are some events and $\mathbb{P}(B) \neq 0$, then

$$\mathbb{P}(A_1 \cup A_2|B) = \mathbb{P}(A_1|B) + \mathbb{P}(A_2|B) - \mathbb{P}(A_1 \cap A_2|B);$$

g. If A is an event with $\mathbb{P}(A) \neq 0$, then

$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B) \cdot \mathbb{P}(B) = \mathbb{P}(B|A) \cdot \mathbb{P}(A);$$

h. (multiplication or chain rule) If $A_1, ..., A_n$ are some events, then

$$\mathbb{P}(A_1 \cap A_2 \cap ... \cap A_n) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2|A_1) \cdot \mathbb{P}(A_3|A_1 \cap A_2) \cdot ...$$

$$\cdot \mathbb{P}(A_n|A_1 \cap A_2 \cap ... \cap A_{n-1}).$$

Ways to use Conditional Probabilities

Conditional Probabilities appear in the following situations:

- We want to calculate $\mathbb{P}(A|B)$. If it is easy to calculate $\mathbb{P}(B)$ and $\mathbb{P}(A \cap B)$, then we calculate $\mathbb{P}(A|B)$ by

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

- We want to calculate $\mathbb{P}(A \cap B)$. If it is easy to calculate $\mathbb{P}(B)$ and $\mathbb{P}(A|B)$, then we calculate $\mathbb{P}(A \cap B)$ by

$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B) \cdot \mathbb{P}(B).$$

This property is called the Multiplication Rule.

The Total Probability Formula

Assume we want to calculate $\mathbb{P}(A)$. We consider Hypotheses B_1, B_2, ..., B_n such that

$$B_i \cap B_j = \varnothing \qquad \text{and} \qquad \bigcup_{k=1}^{n} B_k = \Omega$$

Then

TPF

$$\mathbb{P}(A) = \mathbb{P}(B_1)\cdot\mathbb{P}(A|B_1)+\mathbb{P}(B_2)\cdot\mathbb{P}(A|B_2)+\ldots+\mathbb{P}(B_n)\cdot\mathbb{P}(A|B_n).$$

Tree Form!

Bayes Formula

Bayes Formula

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

We can look at this formula from the following point of view:

- Initially, we have the Probability of B, $\mathbb{P}(B)$; this is the Prior Probability;
- We observe A, new information;
- In the light of this new information, we update our assessment of the Probability of B, we obtain $\mathbb{P}(B|A)$; this is the Posterior Probability.

Bayes Formula

For the general form, assume we have an event A. We have also some Hypotheses B_1, B_2,..., B_n.

- The Direct Problem: Calculate the Probability of A, given the Hypotheses. Solution: TPF

$$\mathbb{P}(A) = \mathbb{P}(B_1)\cdot\mathbb{P}(A|B_1)+\mathbb{P}(B_2)\cdot\mathbb{P}(A|B_2)+...+\mathbb{P}(B_n)\cdot\mathbb{P}(A|B_n)$$

- The Inverse Problem: Calculate the Probability that B_k happened, given that we have observed A. Solution: Bayes Formula:

$$\mathbb{P}(B_k|A) = \frac{\mathbb{P}(A|B_k)\cdot\mathbb{P}(B_k)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B_k)\cdot\mathbb{P}(B_k)}{\sum_{i=1}^{n}\mathbb{P}(A|B_i)\cdot\mathbb{P}(B_i)}.$$

Independence of Events, Definition

Assume we have two Events A and B in the same experiment.

Independence of Events

We say that the Events A and B are **Independent**, if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B).$$

Remark: It is easy to see that the condition above is equivalent, except the cases when $\mathbb{P}(A) = 0$ or $\mathbb{P}(B) = 0$, to

$$\mathbb{P}(A|B) = \mathbb{P}(A)$$

and

$$\mathbb{P}(B|A) = \mathbb{P}(B).$$

Some Properties of Independent Events

Assume A, B are some events.

- If A and B are independent, then A and B cannot be disjoint (unless $\mathbb{P}(A) = 0$ or $\mathbb{P}(B) = 0$);
- If A and B are independent, then so are A and $\overline{B}$, and also $\overline{A}$ and B, also $\overline{A}$ and $\overline{B}$;
- If A is independent of B and A is independent of C, and also $B \cap C = \varnothing$, then A is independent of $B \cup C$.

Independence of more that two events

Now assume we have several events in the Experiment: A_k, $k = 1, 2, ..., n$. We define two notions:

Pairwise Independence

We will say that the events $A_1, A_2, ..., A_n$ are **Pairwise Independent**, if every pair A_i and A_j are Independent, for any $i \neq j$, i.e., if

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i) \cdot \mathbb{P}(A_j), \qquad i \neq j.$$

Independence of more that two events, cont'd

Mutual Independence

We say that $A_1, \dots, A_n$ are **Mutually Independent** or just **Independent**, if for any subgroup of events $A_{i_1}, A_{i_2}, \dots, A_{i_k}$,

$$\mathbb{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = \mathbb{P}(A_{i_1}) \cdot \mathbb{P}(A_{i_2}) \cdot \dots \cdot \mathbb{P}(A_{i_k}),$$

that is, if

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i) \cdot \mathbb{P}(A_j), \qquad i \neq j,$$

$$\mathbb{P}(A_i \cap A_j \cap A_k) = \mathbb{P}(A_i) \cdot \mathbb{P}(A_j) \cdot \mathbb{P}(A_k), \qquad i \neq j, i \neq k, j \neq k,$$

...

$$\mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_n) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2) \cdot \dots \cdot \mathbb{P}(A_n).$$

Independence of more that two events, cont'd

Now, assume we want to calculate the Probability

$$\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_n).$$

- If we know that $A_1, A_2, \ldots, A_n$ are Independent, then

$$\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_n) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2) \cdot \mathbb{P}(A_3) \cdot \ldots \cdot \mathbb{P}(A_n);$$

- In the general case, we have

$$\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_n) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2 | A_1) \cdot \mathbb{P}(A_3 | A_1, A_2) \cdot \ldots \cdot \mathbb{P}(A_n | A_1, \ldots, A_n)$$

Note: Usually, in Probability, the notation $\mathbb{P}(A, B)$ means $\mathbb{P}(A \cap B)$, i.e., say. the above Probability

$$\mathbb{P}(A_3 | A_1, A_2) = p(A_3 | A_1 \cap A_2).$$

Good Luck with MT1!