CS 107, Probability, Spring 2020
Lecture 11

Michael Poghosyan
mpoghosyan@aua.am

AUA

12 February 2020

Content

- Conditional Probabilities
- Total Probability Formula

LZ

The Monty Hall Problem

https://www.youtube.com/watch?v=mhlc7peGlGg

Last Lecture ReCap

Last time we were talking about Conditional Probabilities:

- $\mathbb{P}(A|B) = \dfrac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$;
- $\mathbb{P}(\overline{A}|B) = 1 - \mathbb{P}(A|B)$;
- $\mathbb{P}(A \cap B) = \mathbb{P}(A|B) \cdot \mathbb{P}(B) = \mathbb{P}(B|A) \cdot \mathbb{P}(A)$;
- $\mathbb{P}(A \cap B)$ and $\mathbb{P}(A|B)$ are different things, in general

Interpretation of the Conditional Probability

Last time we were talking about constructing the Conditional Probability Space: Assume we have a Probability Space $(\Omega, \mathcal{F}, \mathbb{P})$. We fix the Condition: $B \subset \Omega$, with $\mathbb{P}(B) \neq 0$. We want to have a Probability Space $(B, \mathcal{F}_B, \mathbb{P}_B)$, where

- B is our new Sample Space;
- $\mathcal{F}_B$ is our new set of Events;
- $\mathbb{P}_B$ is a Probability (Measure) on B.

Interpretation of the Conditional Probability

Now, let us construct $(B, \mathcal{F}_B, \mathbb{P}_B)$:

- our new Sample Space is B;
- for every Event $A \subset \Omega$, our new Event in this new SS is $A \cap B \subset B$; so

$$\mathcal{F}_B = \{A \cap B : A \in \mathcal{F}\};$$

- our new Probability Measure in the new SS is

$$\mathbb{P}_B(\cdot) = \mathbb{P}(\cdot \cap B) \cdot \frac{1}{\mathbb{P}(B)} = \mathbb{P}(\cdot|B),$$

$$\textit{i.e.,} \qquad \mathbb{P}_B(A) = \mathbb{P}(A \cap B) \cdot \frac{1}{\mathbb{P}(B)} = \mathbb{P}(A|B), \qquad \forall A.$$

And, using the above Properties, one can show that $\mathbb{P}_B$ satisfies all axioms of a Probability!

Intermezzo: Recommender Systems

Recommender Systems are widely used in Marketing: say, when you are buying a good from Amazon Store, or watching a movie online, or ordering a food online, then usually you are getting some ads (recommendations) what to buy also, what to watch next, what to order with the chosen qyabab.
The very simple idea behind Recommender Systems is to calculate Conditional Probabilities. Say, we have films A, B, C, D in our webpage, a customer watched already A and B. Which one to recommend: C or D?. The idea is: calculate

$$\mathbb{P}(\text{watch C}|\text{watched A and B}) \quad \textit{and} \quad \mathbb{P}(\text{watch D}|\text{watched A and B}).$$

If the first one is larger, recommend next to watch C, otherwise, recommend D. And to calculate these Conditional Probabilities, we can use the data of all our webpage customers/visitors. This is the part of Statistics and Data Science.

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.

Example:

Problem: We have 28 participants in our Prob class, from which 18 are male. Assume we know that only 20% of our female students love Jazz. We choose a participant from our class at random. What is the Probability that the chosen person will be a Jazz lover female student?

Translation into the *Probability Language*: We want to calculate

$$\mathbb{P}(\text{chosen person is a female and she loves Jazz})$$

We know that

$$\mathbb{P}(\text{chosen person is a female}) =$$

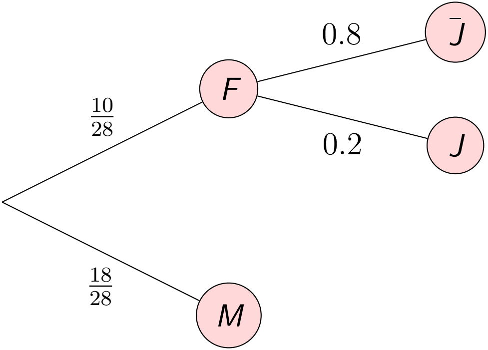
Also the statement **we know that only 20% of our females love Jazz** can be written as:

$$\mathbb{P}(\text{ person loves Jazz}|\text{ person is a female}) = 0.2.$$

Now, $\mathbb{P}(\text{chosen person is a female and she loves Jazz}) = \ldots$

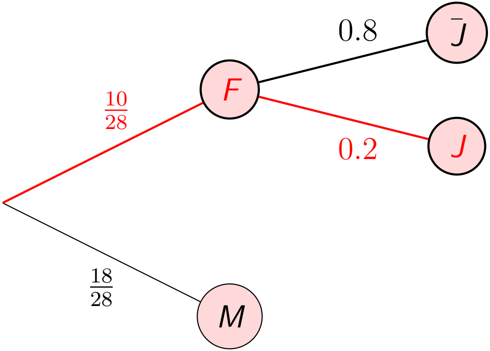
Jazz Problem, Cont'd

The above can be given in the Tree form:



Jazz Problem, Cont'd

The above can be given in the Tree form:



Jazz Problem, Cont'd

Now, let us complete the tree! Recall the above problem: We have 28 participants in our Prob class, from which 18 are male. Assume we know that only 20% of our female students love Jazz. We choose a participant from our class at random. What is the Probability that the chosen person will be a Jazz lover female student?

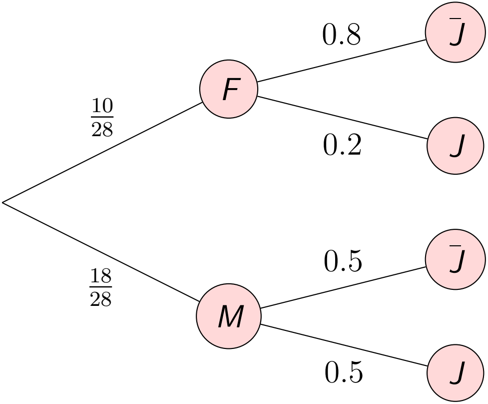
Assume, that in the problem above we know additionally that 50% of male participants of our class love Jazz.

Then we can calculate the Probability that the chosen person will be a Jazz lover male participant.

$$\mathbb{P}(\text{chosen person is a male and he loves Jazz}) =$$

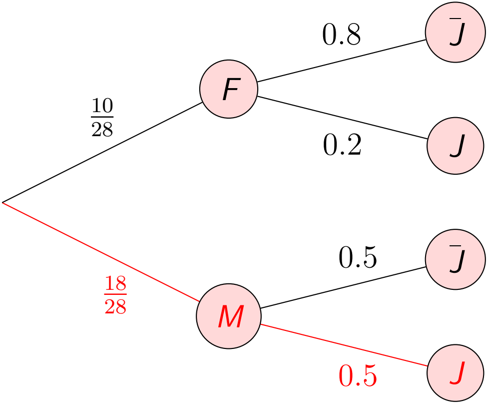
Jazz Problem, Cont'd

The above can be given in the Probability Tree form:



Jazz Problem, Cont'd

The above can be given in the Probability Tree form:



Jazz Problem, Cont'd

Important is that now we can calculate the probability that the chosen person will like Jazz:

$$\mathbb{P}(\text{chosen person likes Jazz}) =$$

$$= \mathbb{P}(\text{chosen person is a male and he loves Jazz}) +$$

$$+ \mathbb{P}(\text{chosen person is a female and she loves Jazz}) =$$

because the two events above (being a male Jazz lover and female jazz lover) are disjoint.

The result can be visualized using the Probability Trees: OTB.

The Total Probability Formula

Now, let us generalize the formula: Assume A and B are events in some Experiment. Then

Total Probability Formula

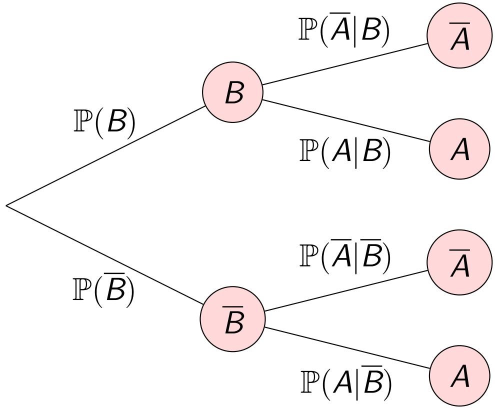
$$\mathbb{P}(A) = \mathbb{P}(B) \cdot \mathbb{P}(A|B) + \mathbb{P}(\overline{B}) \cdot \mathbb{P}(A|\overline{B})$$

Here B and $\overline{B}$ are called Hypotheses, and satisfy

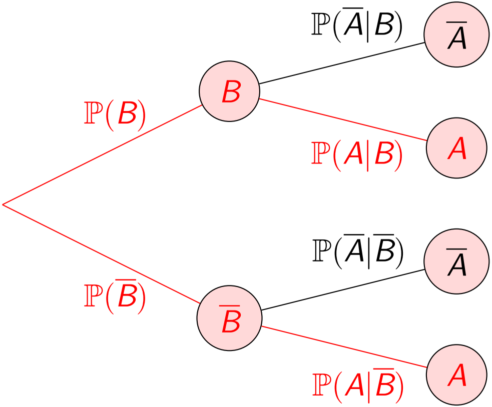
$$B \cap \overline{B} = \varnothing \qquad \text{and} \qquad B \cup \overline{B} = \Omega,$$

that is, **we take a partition of our Sample Space into disjoint Hypotheses (Events), calculate the probability of our event A under each Hypothesis, and then add the results**.

TPF, Probability Tree form



TPF, Probability Tree form



The Total Probability Formula

The general form of the TPF is: 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

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

Example

Problem: We have an urn full of 100 balls, from which 15 are white, 25 are red, and the rest are black. Our balls are of two type: round and cubic $\ddot\smile$. The proportion of round balls in whites is 20%, in reds - 40%, and in blacks - 55%. We take a ball at random. What is the probability of drawing a cubic ball?

Well, Probabilists love urns and balls, and dice and coins. The problem may seem very artificial. So let me give a problem of this type in a more realistic terms:

Example

Problem: We know that in some city, 15% of persons are below the age 15, 25% are between 15 and 30, and the rest of population in that city are above 30. The proportion of persons in that city, regularly visiting gyms, for the young-age population is 20%, for mid-age-population - 40%, and for older-age-population - 55%. What is the probability of that a person in that city will be a cubic ball? I mean, the person will be a regular gym visitor? $\ddot\smile$

Solution: OTB

Example

Problem: At our YSU, we are usually doing oral exams. Students are randomly taking exam papers (exam papers differ from each other, each exam paper can contain 2-3 theoretical questions), preparing their answers for some time, and then presenting them to the lecturer. Assume our YSU Prob Lecturer prepared 27 exam papers. 5 of these papers are easy ones, and others are veeer... not-so easy ones. Now, two students are entering the exam room to take the papers. Whose chances are higher to take an easy exam paper: the first or the second student?

Solution: OTB

Bayes Formula

Now, let us talk about a very important formulas, the Bayes Formula. It is used a lot in Statistics, Machine Learning, Signal Processing, Finance, Bioinformatics etc.
The formula is simple: we just write the property of the Conditional Probabilities:

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

Hence,

Bayes Formula

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

Bayes Formula, Cont'd

In the Bayes formula,

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

- $\mathbb{P}(B)$ is the **Prior Probability** of B
- $\mathbb{P}(B|A)$ is the **Posterior Probability** of B, after observing A
- $\mathbb{P}(A|B)$ is called the **Likelihood** of A under B.

Bayes Formula, Interpretation

To explain the above names, let us write again the Bayes Formula:

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

Now, 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.