Probability 2 : Conditional Probability

Anand Systla

Masters in Financial Engineering Bootcamp
UCLA Anderson

August 10, 2022

TODAY'S AGENDA

- Discuss the previous class take home questions

- Independence

- Conditional probability

- Baye's Theorem + Example

- Conditional independence

- Simpson's paradox

- Random Variables

Conditional Probability Motivation

- How do we change our probability of an event occuring given new information?

- Say we want the probability of it snowing tomorrow and we're given
  1. You are located in Antarctica
  2. It is the month of December
  3. Humidity today is 80%

- A really neat way to think about this is using a probability tree diagram

CONDITIONAL PROBABILITY

- Conditional probability $A|B$ means we condition on a new information set B.

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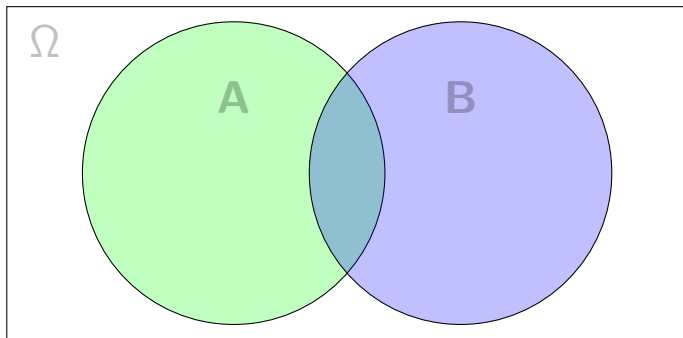
INDEPENDENCE

- Two events A and B are independent iff $P(A \cap B) = P(AB) = P(A) \cdot P(B)$

- We can show independence in two ways

    ▶ Assumption: Two coin tosses, coin toss and the color of the building you're conducting the experiment, etc.
    ▶ Derivation: $P(AB) = P(A) \cdot P(B)$

- **Example:** Say $\Omega = \{1, 2, 3, 4, 5, 6\}$, $A = \{1, 2, 3, 4\}$, $B = \{2, 4, 6\}$, $C = \{1, 6\}$

    ▶ $P(A) = \hspace{3cm} P(B) = \hspace{3cm} P(C) =$

    ▶ $P(AB) = P(A \cap B = \{2, 4\}) =$

    ▶ $P(A) \cdot P(B) =$

    ▶ $P(AC) = P(A \cap C = \{1\}) =$

    ▶ $P(A) \cdot P(C) =$

    ▶ Alternatively, if B has occurred (we focus on $\{2, 4, 6\}$), then $P(A|B = \{2, 4\}) =$

    ▶ Occurrence of event B has no effect on the likelihood of occurrence of event A (what about C?)

- If A and B are independent, how does $P(A)$ change given B has occurred? i.e. $P(A|B) =$

Independent vs Disjoint Events

- In a single coin toss, is the heads outcome independent of the tails outcome?

Ω

- Examples of disjoint events: sides of a dice, positive/negative returns on a day
- Examples of independent events: Two consecutive coin tosses, yearly stock returns (maybe?)

Bayes Theorem

- One of the most important theorems in probability. Rearranging the conditional probability expression, we get

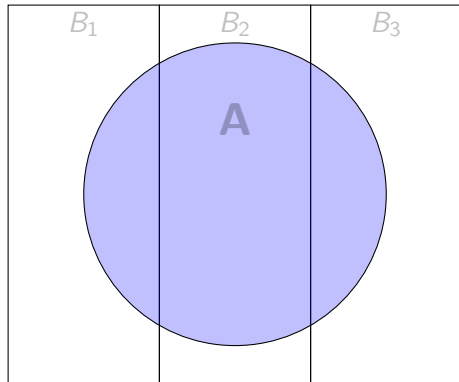$$P(AB) = P(A|B) \cdot P(B)$$
$$P(AB) = P(B|A) \cdot P(A)$$
$$\implies P(A|B) = P(B|A) \cdot \frac{P(A)}{P(B)}$$

- $P(A|B) \neq P(B|A)$
- $P(A|B) \neq 1 - P(A|B^c)$
- $P(A|B) = 1 - P(A^c|B)$
- $P(A|B) = \frac{P(AB)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A)$ if $A \& B$ are independent
- Intuitively,

$$P(\text{Hypothesis}|\text{Data}) = P(\text{Data}|\text{Hypothesis}) \cdot P(\text{Hypothesis})/P(\text{Data})$$
$$P(\text{Flu}|\text{Symptoms}) = P(\text{Symptoms}|\text{Flu}) \cdot P(\text{Flu})/P(\text{Symptoms})$$

Law of Total Probability

- Generalizing $P(A) = \sum P(AB_i) = \sum P(A|B_i)P(B_i)$. Assumptions?
- Applications: If computing $P(A)$ is a challenge (defective parts sold by an industry)
- How can we write the law if we have two general sets A and B?
- How can we combine the law of total probability and Baye's Theorem?

CONDITIONAL INDEPENDENCE

- A and B are conditionally independent given event C iff

$$P(AB|C) = P(A|C) \cdot P(B|C)$$

- What is the relationship between unconditional and conditional independence?

- Three situations we will study
  - Unconditionally independent and conditionally independent
  - Unconditionally independent, but conditionally dependent
  - Unconditionally dependent, but conditionally independent

Unconditionally Independent, Conditionally Dependent

- $\Omega = \{1, 2, 3, 4, 5, 6\}, A = \{1, 2\}, B = \{2, 4, 6\}, C = \{1, 2, 3, 4, 5\}$

- $P(A) = \qquad P(B) = \qquad P(C) = \qquad P(AB) =$

- Are A and B unconditionally independent?

- $P(A|C) = \qquad P(B|C) = \qquad P(AB|C) =$

- Conditional on event C, are A and B conditionally independent?

- What would a real life example look like?

UNCONDITIONALLY DEPENDENT, CONDITIONALLY INDEPENDENT

- $\Omega = \{1, 2, 3, 4, 5, 6\}, A = \{1, 2, 5\}, B = \{2, 4, 6\}, C = \{1, 2, 3, 4\}$

- $P(A) = \qquad P(B) = \qquad P(C) = \qquad P(AB) =$

- Are A and B unconditionally independent?

- $P(A|C) = \qquad P(B|C) = \qquad P(AB|C) =$

- Conditional on event C, are A and B conditionally independent?

- What would a real life example look like?

UNCONDITIONALLY INDEPENDENT, CONDITIONALLY INDEPENDENT

- $\Omega = \{1, 2, 3, 4, 5, 6\}, A = \{1, 2\}, B = \{2, 4, 6\}, C = \{1, 2, 3, 4\}$

- $P(A) = \qquad P(B) = \qquad P(C) = \qquad P(AB) =$

- Are A and B unconditionally independent?

- $P(A|C) = \qquad P(B|C) = \qquad P(AB|C) =$

- Conditional on event C, are A and B conditionally independent?

- What would a real life example look like?

Simpson's Paradox

- Classic example of how conditioning and aggregation affect things. Prof. Joe Blitzstein calls conditioning the "soul of statistics" and that all probabilities are conditional.

- Whenever we have data, we want to think about how it was sampled. We need to be careful what we are conditioning on.

Bayes Theorem Example

- Probability of a car theft in a neighbourhood is 1%. All cars are now equipped with state-of-the-art alarms that have a success rate of 90% (will go off 90% of the times if there is a theft). However, the alarms sometimes go off randomly (due to a bird, ball etc.) around 20% of the times. What is the probability that there is a theft happening conditional on the alarm going off?

Bayes Theorem Example

- Probability of a car theft in a neighbourhood is 1%. All cars are now equipped with state-of-the-art alarms that have a success rate of 90% (will go off 90% of the times if there is a theft). However, the alarms sometimes go off randomly (due to a bird, ball etc.) around 20% of the times. What is the probability that there is a theft happening conditional on the alarm going off?

$$P(\text{Theft}) = 0.01$$

$$P(\text{Alarm}|\text{Theft}) = 0.9$$

$$P(\text{Alarm}|\text{No Theft}) = 0.2$$

$$P(\text{Alarm}) = P(\text{Alarm}|\text{Theft}) \cdot P(\text{Theft}) + P(\text{Alarm}|\text{No Theft}) \cdot P(\text{No Theft})$$

$$P(\text{Alarm}) = 0.9 \times 0.01 + 0.2 \times 0.99 = 0.009 + 0.198 = 0.207$$

$$\implies P(\text{Theft}|\text{Alarm}) = P(\text{Alarm}|\text{Theft}) \cdot \frac{P(\text{Theft})}{P(\text{Alarm})} = 0.9 \cdot \frac{0.01}{0.207} = 0.043$$

TAKE HOME QUESTIONS

1. **Quant Interview:** A quant driven hedge fund wants to interview all the UCLA MFE students for an internship. Say 50% of all students who received their first interview received a second interview. 95% of your friends that got a second interview said they had a good first interview. 75% of your friends that did not get a second interview said they had a good first interview. If you felt you had a good first interview, what is the probability that you will receive a second interview? Alternatively, if you felt you had a bad first interview, what is the probability that you will receive a second interview?

2. There are two biased coins A and B in a bag. Probability of heads for coin A is 0.75 and the probability of heads for coin B is 0.3. You pick a coin randomly and perform 10 tosses (without knowing which coin you picked). Hint: To solve the two problems below, compute the posterior probability $P(Hypothesis|Data)$ and argue that one of the coin has a higher posterior probability. You will have to test and compare the two hypothesis - picking coin A given data and picking coin B given data. Since we are choosing a coin randomly, P(picking coin A) = P(picking coin B) = 1/2. Key takeaway is how the data changes your beliefs about which coin you picked.

   2.1 You observe that you get 8 heads and 2 tails from your coin tosses. What is the probability that you picked coin A from the bag given the data. Compare this with the posterior probability of picking the other coin.

   2.2 Now, say you observed 8 tails and 2 heads from your coin tosses. What is the probability that you picked coin B given the data. Compare this with the posterior probability of picking the other coin.

1. Goal was to get you to use Baye's Theorem. Also, $P(A|B) \neq P(A|B^c)$

   1.1 Defining events as {Good,Bad} and {Pass,Fail}

   $$P(\text{Pass}) = P(\text{Fail}) = 0.5$$

   $$P(\text{Good}|\text{Pass}) = 0.95$$

   $$P(\text{Good}|\text{Fail}) = 0.75$$

   $$P(\text{Good}) = P(\text{Good}|\text{Pass}) \cdot P(\text{Pass}) + P(\text{Good}|\text{Fail}) \cdot P(\text{Fail}) = (0.95)(0.5) + (0.75)(0.5) = 0.85$$

   $$P(\text{Pass}|\text{Good}) = P(\text{Good}|\text{Pass}) \cdot \frac{P(\text{Pass})}{P(\text{Good})} = 0.95 \cdot \frac{0.5}{0.85} = 0.5588$$

   1.2 Now considering the alternate situation -

   $$P(\text{Bad}|\text{Pass}) = 0.05$$

   $$P(\text{Bad}|\text{Fail}) = 0.25$$

   $$P(\text{Bad}) = P(\text{Bad}|\text{Pass}) \cdot P(\text{Pass}) + P(\text{Bad}|\text{Fail}) \cdot P(\text{Fail}) = 0.5 * (0.05 + 0.25) = 0.15$$

   $$P(\text{Pass}|\text{Bad}) = P(\text{Bad}|\text{Pass}) \cdot \frac{P(\text{Pass})}{P(\text{Bad})} = 0.05 \cdot \frac{0.5}{0.15} = 0.167$$

2. Goal of this exercise was to show you that, by changing the data, your belief (posterior) about the outcome changes! Note : Our prior is uniform (more later).

    2.1 8 Heads and 2 Tails:

$$P(\text{coin A}) = P(\text{coin B}) = 1/2$$

$$P(\text{Data}|\text{coin A}) = {}^{10}C_2(0.75)^8(0.25)^2 = {}^{10}C_2 \times 0.00626$$

$$P(\text{Data}|\text{coin B}) = {}^{10}C_2(0.3)^8(0.7)^2 = {}^{10}C_2 \times 0.00003$$

$$P(\text{Data}) = P(\text{Data}|\text{coin A})P(\text{coin A}) + P(\text{Data}|\text{coin B})P(\text{coin B})$$

$$P(\text{Data}) = {}^{10}C_2\frac{1}{2} \cdot 0.00626 + {}^{10}C_2\frac{1}{2} \cdot 0.00003 = {}^{10}C_2 \times 0.00314$$

$$P(\text{coin A}|\text{Data}) = P(\text{Data}|\text{Coin A}) \cdot \frac{P(\text{coin A})}{P(\text{Data})} = 0.00626 \cdot \frac{0.5}{0.00314} \cdot \frac{{}^{10}C_2}{{}^{10}C_2} = 0.996$$

$$P(\text{coin B}|\text{Data}) = P(\text{Data}|\text{Coin B}) \cdot \frac{P(\text{coin B})}{P(\text{Data})} = 0.00003 \cdot \frac{0.5}{0.00314} \cdot \frac{{}^{10}C_2}{{}^{10}C_2} = 0.004$$