MA-120 Probability and Statistics

Nazar Khan

PUCIT

Lecture 9: Conditional Probability

CONDITIONAL PROBABILITY

Motivation

- So far, we have studied the probability of events.
  - P(Pakistan winning a match)
- What if we have some background information
  - P(Pakistan winning a match knowing that the main player is injured)
- Such questions are answered through **conditional probability**.

Conditional Probability

Let sample space S be all points in the green rectangle.
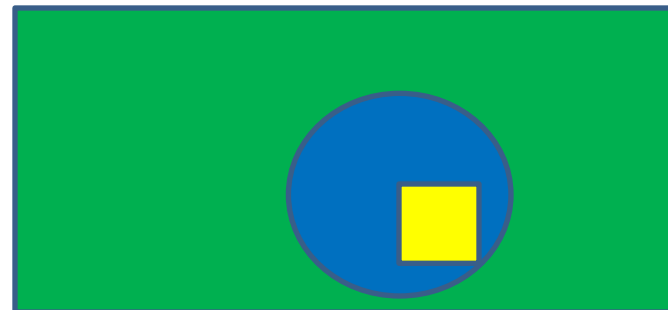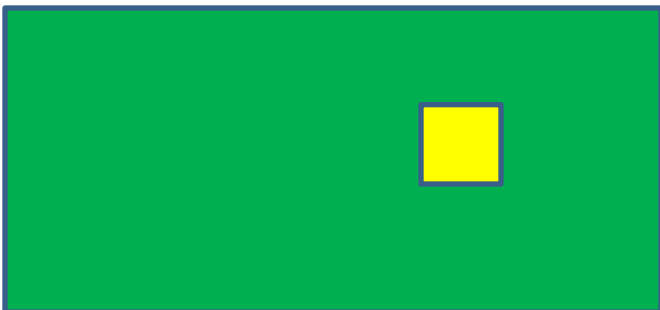Let event A be any point in the yellow square
Let event B be any point in the blue circle.

Probability

- P(point in yellow square) = area(yellow square)/area(S)

- P(A)=area(A)/area(S)

Conditional Probability

- P(point in yellow square | point is in blue circle) = area(yellow square) / area(blue circle)

- P(A|B) = area(A∩B)/area(B) = area(A∩B)/area(S) / area(B)/area(S)

  = P(A∩B)/P(B)

Conditional Probability

- P(A|B)=area(A∩B)/area(B)

- Notice that the denominator contains area(B) instead of area(S).

- So event B is acting as a **reduced sample space** or a **sub-sample space**.

- The basic idea of conditional probability is that the **conditioning event serves as a sub-sample space**.
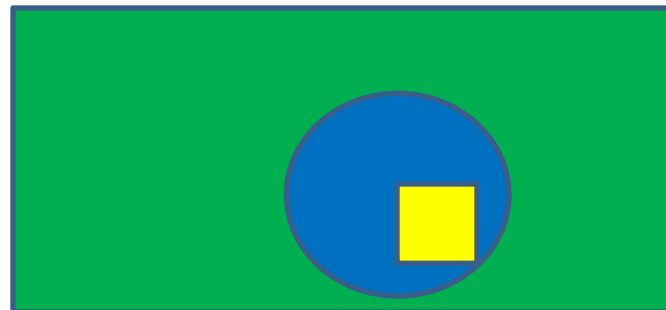
Conditional Probability

- Probability
  - P(wet ground)
  - P(it's raining)
- Conditional Probability
  - **P(A|B)=P(A∩B)/P(B)**
  - P(wet ground | it's raining) = P(W∩R)/P(R)
  - P(it's raining | wet ground) = P(R∩W)/P(W)

Conditional Probability

Example 1

- Suppose that out of all fans in Punjab, 20% are manufactured in Lahore, 50% in Gujrat and 30% in Faisalabad.
- The %age of defective fans in the 3 cities are 5%, 3% and 4% respectively.
- Let D be the event that a fan is defective.
- Let L be the event that it was manufactured in Lahore, similarly F for Faisalabad and G for Gujrat.
- **What is P(L), P(F) and P(G)?**
- **What is P(D|G)?**

Conditional Probability

- P(L)=P(fan was manufactured in Lahore)

  $$= \frac{|\text{fans manufactured in Lahore}|}{|\text{fans manufactured in Punjab}|}$$

  = ratio of Lahore fans = 20%.

- Similarly, P(G)=50% and P(F)=30%.

Conditional Probability

- P(D|G)

  =P(fan was defective **given** that it was manufactured in Gujrat)

  = $\dfrac{|\text{defective fans manufactured in Gujrat}|}{|\text{fans manufactured in Gujrat}|}$

  = ratio of defective Gujrat fans = 3%.

Conditional Probability

Example 2

- We draw a card at random from the standard deck of 52 playing cards.
- A = {card drawn is ace of hearts}
- B = {card drawn is red }.
- **What is P(A|B)?**
  - Convert to a sentence first.
  - **You cannot solve what you don't understand!**
  - What is the reduced sample space?

Key:
♣=club (13 in total)
♦=diamond (13 in total)
♥=heart (13 in total)
♠=spade (13 in total)

Conditional Probability

- Using reduced sample space.
  - There are only 26 red cards and there is only 1 ace of hearts in them. Therefore, P(A|B)=1/26
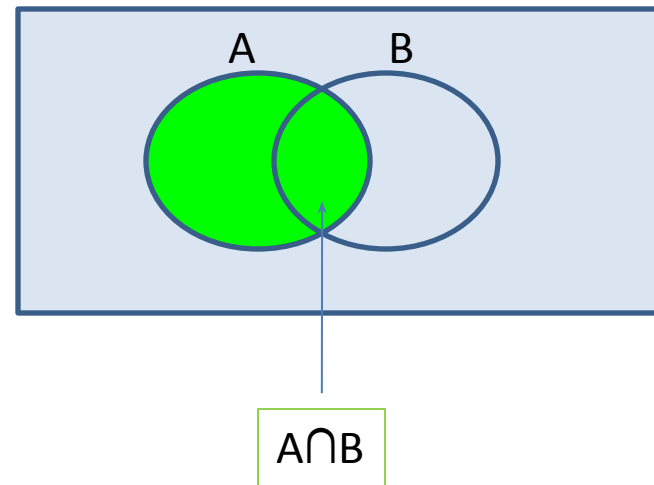- Using the formula.
  - P(A|B) = P(A and B)/P(B) = (1/52) / (1/2) = 1/26

Properties

$P(A^c|B)$

$=P(A^c \cap B)/P(B)$

$=(P(B)-P(A \cap B)) / P(B)$

$=1-P(A|B)$

Properties

For disjoint events A and C,

$P(A \cup C | B)$

$= P((A \cup C) \cap B)/P(B)$

$= P(((A \cap B) \cup (C \cap B)))/P(B)$

$= (P(A \cap B) + P(C \cap B)) / P(B)$

$= P(A \cap B)/P(B) + P(C \cap B)/P(B)$

$= P(A|B) + P(C|B)$
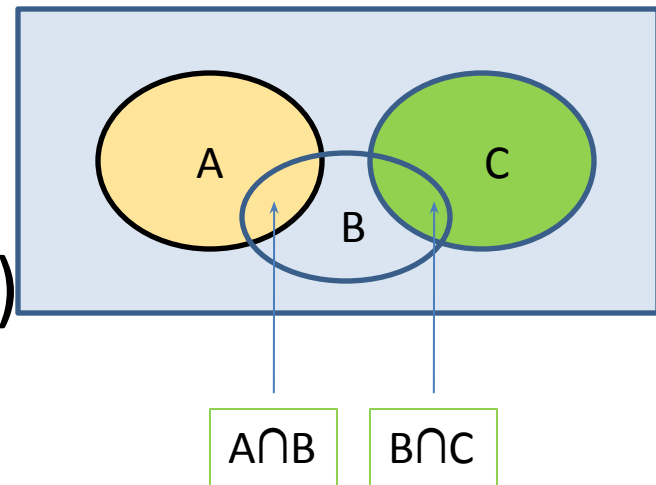
Properties

- $P(A|B) = P(AB)/P(B) \Rightarrow P(AB) = P(A|B)P(B)$
- $P(B|A) = P(AB)/P(A) \Rightarrow P(AB) = P(B|A)P(A)$
- So we can write
  $P(ABC)$
  $= P(A|BC)P(BC)$
  $= P(A|BC)P(B|C)P(C)$
- We can also write
  $P(ABC)$
  $= P(C|AB)P(AB)$
  $= P(C|AB)P(B|A)P(A)$

Properties

- P(A|B)=P(AB)/P(B)
- What if A and B are independent?
  - P(AB)=P(A)P(B)
  - So P(A|B)=P(A)
  - **Similarly P(B|A)=?**
- What if A and B are mutually exclusive?
  - P(AB)=0
  - So P(A|B)=0
  - P(B|A)=0

Properties

Probability

- $P(S)=1$

- $P(A^c)=1-P(A)$

- $P(A \cup C)=P(A)+P(C)$

- For independent A,B
$P(AB)=P(A)P(B)$

Conditional Probability

- $P(S|B)=1$

- $P(A^c|B)=1-P(A|B)$

- $P(A \cup C|B)=P(A|B)+P(C|B)$

- For independent A,B
  - $P(A|B)=P(A)$
  - $P(B|A)=P(B)$

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

Conditional probability behaves just like ordinary probability with the conditioning event, B, staying along.

Conditional Probability

	P(Good GPA)	P(Bad GPA)	Total
P(Athlete)	.05	.15	.2
P(Non-athlete)	.35	.45	.8
Total	.4	.6	1

$P(A)=P(Athlete)=.2$ $P(A^c)=P(Non\text{-}athlete)=.8$

$P(G)=P(Good\ GPA)=.4$ $P(G^c)=P(Bad\ GPA)=.6$

$P(A \cap G)=.05$ $P(A \cup G)=.55$

$P(A|G)=.05/.4$ $P(G|A)=.05/.2$

$P(A \cup G | A \cap G)=1$ $P(A \cap G | A \cup G)=.05/.55$

THEOREM OF TOTAL PROBABILITY

Theorem of Total Probability

Example 1 Revisited

- Suppose that out of all fans in Punjab, 20% are manufactured in Lahore, 50% in Gujrat and 30% in Faisalabad.
- The %age of defective fans in the 3 cities are 5%, 3% and 4% respectively.
- Let D be the event that a fan is defective.
- Let L be the event that it was manufactured in Lahore, similarly F for Faisalabad and G for Gujrat.
- Earlier, we found P(L), P(F), P(G) and P(D|G)?
- **What is P(D)?**

Theorem of Total Probability

Example 1 Revisited

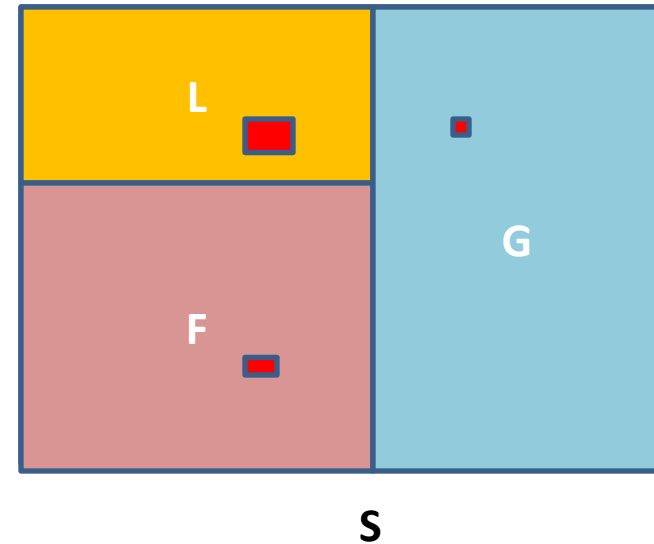- P(L)=.2, P(G)=.5, and P(F)=.3.
- P(D|L)=.05, P(D|G)=.03, P(D|L)=.04
- P(D)

=P(D∩L)+P(D∩G)+P(D∩F)

=P(D|L)P(L)+P(D|G)P(G)+ P(D|F)P(F)

= (0.05 x .20) + (0.03 x .50) + (0.04 x .30)

=.037

S

Theorem of Total Probability

Theorem of Total Probability: If events $B_1, B_2, \ldots$ form a partition of the sample space S, then for any event A,

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \ldots$$

Proof:

R.H.S $\square$ $P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \ldots$

$= P(AB_1)/P(B_1)*P(B_1) + P(AB_2)/P(B_2)*P(B_2) + \ldots$

$= P(AB_1) + P(AB_2) + \ldots$

$= P(AS) = P(A) \square$ L.H.S

Bayes Theorem

- **P(B|A)**
  =P(AB)/P(A)
  =P(AB)/P(A) * P(B)/P(B)
  =**P(A|B)P(B)/P(A)** since P(AB)/P(B)=P(A|B)
- **This is known as Bayes Theorem**.
- To find the most likely reduced sample space from which a given outcome could have come
  - P(G|D)=P(Fan was manufactured in Gujrat given that it was defective)

Bayes Theorem

- P(G|D)

  $= P(DG)/P(D)$

  $= P(D|G)P(G)/P(D)$

  $= (.03*.5)/.037 = .405$

Combining Everything Together

Bayes Theorem

$P(B_k | A) = P(A | B_k)P(B_k) / P(A)$

$\qquad = P(A | B_k)P(B_k) / \sum_i P(A | B_i)P(B_i)$

Practice Questions

1. Prove or disprove, $P(A|B^c) = 1 - P(A|B)$ when $0 < P(B) < 1$.

2. Prove or disprove, $P(A|B \cup C) = P(A|B) + P(A|C) - P(A|B \cap C)$, when $P(B \cap C) > 0$.

3. Find $P(A|B)$ when $A \subseteq B$ and $P(B) > 0$.

4. Let A,B have positive probabilities. If A,B are independent then what are the values of $P(A|B)$ and $P(B|A)$?

5. Let A,B have positive probabilities. If $P(A) \leq P(A|B)$ then prove that $P(B) \leq P(B|A)$.

6. If A,B are independent then find $P(A^c|B^c)$.

7. When A,B are disjoint events, with $P(B) > 0$, what are $P(A|B)$ and $P(A|A \cup B)$?

Practice Questions

8. **There are three boxes, A, B, C, such that box A has 10 red and 16 green balls, box B has 11 red and 15 green balls, and box C has 12 red and 14 green balls. A box is selected at random and from the selected box a ball is drawn at random. What is the probability that the drawn ball is red?**

Practice Questions

9. **Three branches (factories) of a firm, A,B,C, respectively produce 50%, 40% and 10% of the total output of the firm. The percentages of the defective output of these factories are respectively 3%, 2% and 1%. An item is selected at random from the pooled output of the three factories and is found to be defective. What is the probability that it was manufactured by factory A?**

Practice Questions

10. **USB Flash-drives were imported from four countries, C1,C2,C3, C4. The percentage of the total amount imported from the respective countries were 30%, 40%, 20%, 10%. The drives were then packed into small identical looking packages of one hundred drives each. (Each package contained drives from only one country). Certain proportions of the drives were faulty. The percentage of faulty drives from the respective four countries were 3%, 2%, 1%, 1%. Out of one package, five drives were checked and 2 of them turned out to be faulty. What is the probability that the package came from country C1?**

Practice Questions

11. **A box contains two green and four red balls. Two balls are randomly drawn one after the other without replacement.**

    a) **What is the probability that the second ball is red?**

    b) **What is the probability that the second ball is red given that the first ball was red?**

    c) **What is the probability that the first ball was red given that the second ball drawn was red?**