

Probability Theory

Chapter 3—Conditional Probability and Independence

Nan Liu¹

¹These slides are based on Prof. Mui Li's. Any mistakes are my own. If you find an error, please send an email at nanliu@xmu.edu.cn

3.2 Conditional Probabilities

- **Conditional Probability.** Consider an experiment with sample space S . Let E and F be two events, then the conditional probability of E given F is denoted by $P(E|F)$ and satisfies if $P(F) > 0$

$$P(E|F) = \frac{P(E \cap F)}{P(F)}$$

- **Intuition:** If F has occurred, then, in order for E to occur, it is necessary that the occurrence be both in E and F , hence it must be in $E \cap F$. Once F has occurred, F is the new sample space.
- *Equally likely outcomes.* In this case, we have

$$\begin{aligned} P(E|F) &= \frac{\# \text{ outcomes in } E \cap F}{\# \text{ outcomes in } F} \\ &= \underbrace{\frac{\# \text{ outcomes in } E \cap F}{\# \text{ outcomes in } S}}_{P(E \cap F)} / \underbrace{\left(\frac{\# \text{ outcomes in } F}{\# \text{ outcomes in } S} \right)}_{P(F)}. \end{aligned}$$

3.2 Example

- Mr and Mrs. Smith have 2 children, boys or girls, the 4 configurations being equally likely. What is the probability that the two children are two girls if **1)** you are not given any additional information, **2)** the elder child is a girl, **3)** one of the two children is a girl.

- **Answer to 1):** By assumption, we have

$P(\{G, G\}) = P(\{G, B\}) = P(\{B, G\}) = P(\{B, B\}) = \frac{1}{4}$ so if you don't have any additional information $P(\{G, G\}) = 1/4$.

- **Answer to 2):** Now we know that the elder child is a girl so the event $F = (\{G, G\}, \{G, B\})$ occurred. We are interested in

$$P(\{G, G\} | F) = \frac{P(F \cap \{G, G\})}{P(F)} = \frac{P(\{G, G\})}{P(F)} = \frac{\frac{1}{4}}{\frac{2}{4}} = \frac{1}{2}.$$

- **Answer to 3):** Now we know that one of the two children is a girl so the event $F = (\{G, G\}, \{G, B\}, \{B, G\})$ occurred. We have

$$P(\{G, G\} | F) = \frac{P(F \cap \{G, G\})}{P(F)} = \frac{P(\{G, G\})}{P(F)} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}.$$

3.2 Examples

- **Example:** A math teacher gave her class two tests. 25% of the class passed both tests and 42% of the class passed the first test. What percent of those who passed the first test also passed the second test?
- **Answer:** Let E_i = “Event passed test i ”. We have

$$P(E_1) = 0.42 \text{ and } P(E_1 \cap E_2) = 0.25$$

so

$$P(E_2|E_1) = \frac{P(E_1 \cap E_2)}{P(E_1)} = \frac{0.25}{0.42} = 0.60$$

- **Example:** The probability that it is Friday and that a student is absent is 0.03. Since there are 5 school days in a week, the probability that a school day is Friday is 0.2. What is the probability that a student is absent given that today is Friday?
- **Answer:** We have

$$P(\text{Absent}|\text{Friday}) = \frac{P(\text{Absent on Friday})}{P(\text{Friday})} = \frac{0.03}{0.2} = 0.15.$$

3.2 The Multiplication Rule

- Let $E_1, E_2, \dots, E_n$ be a sequence of events, then we have

$$P(E_1 \cap E_2 \cap \dots \cap E_n) = P(E_1) P(E_2 | E_1) \times \\ \times P(E_3 | E_1 \cap E_2) \dots P(E_n | E_1 \cap \dots \cap E_{n-1})$$

- Proof follows from

$$P(E_1 \cap \dots \cap E_n) = \frac{P(E_1 \cap \dots \cap E_n)}{P(E_1 \cap \dots \cap E_{n-1})} P(E_1 \cap E_2 \cap \dots \cap E_{n-1}) \\ = P(E_n | E_1 \cap \dots \cap E_{n-1}) \frac{P(E_1 \cap \dots \cap E_{n-1})}{P(E_1 \cap \dots \cap E_{n-2})} P(E_1 \cap \dots \cap E_{n-2})$$

i.e. apply repeatedly $P(A|B) = P(A \cap B) / P(B)$ where $A = E_1 \cap \dots \cap E_i$ and $B = E_1 \cap \dots \cap E_{i-1}$.

3.2 Example

- **Example:** You have a box with 3 blue marbles, 2 red marbles, and 4 yellow marbles. You are going to pull out one marble, record its color, put it back in the box and draw another marble. What is the proba. of pulling out a red followed by a blue?
- **Answer:** We have $P(\text{red}) = 2/9$ and $P(\text{blue}|\text{red first}) = P(\text{blue}) = 3/9$ so

$$P(\text{red} \cap \text{blue}) = P(\text{red}) P(\text{blue}|\text{red first}) = \frac{2}{9} \times \frac{3}{9} = \frac{2}{27}.$$

- **Example:** Consider the same box of marbles. However, we are going to pull out the first marble, leave it out and then pull out the second marble. What is the probability of pulling out a red marble followed by a blue marble?
- **Answer:** We have $P(\text{red}) = 2/9$ and $P(\text{blue}|\text{red first}) = 3/8 \neq P(\text{blue})$ so

$$P(\text{red} \cap \text{blue}) = P(\text{red}) P(\text{blue}|\text{red first}) = \frac{2}{9} \times \frac{3}{8} = \frac{1}{12}.$$

3.3 Bayes' formula

- **Total probability formula:** We have for any events E, F

$$\begin{aligned}P(E) &= P(E \cap F) + P(E \cap F^c) \\&= P(E|F)P(F) + P(E|F^c)P(F^c) \\&= P(E|F)P(F) + P(E|F^c)(1 - P(F))\end{aligned}$$

- This can be generalized to the following case. Assume $\cup_{i=1}^n F_i = S$ and $F_i \cap F_j = \emptyset$ then

$$\begin{aligned}P(E) &= P(E \cap (\cup_{i=1}^n F_i)) = P(\cup_{i=1}^n (E \cap F_i)) \\&= \sum_{i=1}^n P(E \cap F_i) = \sum_{i=1}^n P(E|F_i)P(F_i)\end{aligned}$$

- **Bayes' formula:**

$$P(F_i|E) = \frac{P(E|F_i)P(F_i)}{\sum_{i=1}^n P(E|F_i)P(F_i)}$$

3.3 Example: Airport Scanner

- A government agency has developed a scanner which determines whether a person is a terrorist. The scanner is fairly reliable; 99% of all scanned terrorists are identified as terrorists, and 95% of all upstanding citizens are identified as such. An informant tells the agency that exactly one passenger of 100 aboard an aeroplane in which in you are seated is a terrorist. The agency decide to scan each passenger, and the shifty looking man sitting next to you tests positive. What are the chances that this man really is a terrorist?
- **Answer.** $A = \text{"Your neighbour is a terrorist"}$, $B = \text{"Your neighbour tested positive"}$. We are interested in $P(A|B)$ where

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}$$

where $P(A) = \frac{1}{100} = 1 - P(A^c)$, $P(B|A) = 0.99$,
 $P(B|A^c) = 1 - 0.95 = 0.05$ so

$$P(A|B) = \frac{0.99/100}{0.99/100 + 0.05 \times 99/100} = 0.1667$$

3.3 Example: Screening for HIV test

- HIV blood tests are very accurate: around 99.8% of people with HIV test positive, and 99.99% of people without the virus test negative. In the UK, the prevalence of HIV in adults with no risk factors is around 1 in 10000. For an adult with no risk factors, what is the probability of having HIV given a positive test result? what is the probability of having HIV given a negative test result?
- **Answer.** Consider the event A = "Adult has HIV" and B = "HIV test positive". We are interested in $P(A|B)$ where

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}.$$

Here we have

$$P(A) = \frac{1}{10000} = 1 - P(A^c), \quad P(B|A) = 0.998, \quad P(B|A^c) = 0.0001$$

so

$$P(A|B) = \frac{0.998/10000}{0.998/10000 + 0.0001 \times 9999/10000} \approx 0.5$$

3.3 Example: Screening for HIV test (cont.)

- The probability of having HIV given a negative test result is $P(A|B^c)$ where

$$P(A|B^c) = \frac{P(B^c|A)P(A)}{P(B^c|A)P(A) + P(B^c|A^c)P(A^c)}$$

where

$$P(B^c|A) = 1 - 0.998 = 0.002, \quad P(B^c|A^c) = 0.9999$$

so

$$P(A|B^c) = \frac{0.002/10000}{0.002/10000 + 0.9999 \times 9999/10000} = 2.10^{-7};$$

there is $2.10^{-5}\%$ chance that the person is HIV positive whereas he/she tested negative.

3.3 Example: Screening for Breast Cancer

- Since 1988, women over 50 in the UK have routinely been offered screening for breast cancer, even if they have no other symptoms, and in 2004/05 1.7 million women were screened. Those with a positive mammogram are recalled for further investigations, at considerable cost in anxiety, resources and pain and discomfort. It has been estimated that in women between 50 and 70 years old, mammography will detect approximately 85% of breast cancers. It has been estimated that around 10% of women with no cancer will still receive a positive result. Finally, we need to know something about the underlying incidence of breast cancer in this population. This varies quite a lot by ethnicity and nationality, but in the UK it has been estimated that around 10 out of 1000 women aged 50 to 70 have breast cancer.
- **Question:** For a woman between 50 and 70 years randomly screened, what is the probability of having breast cancer if she tests positive?

3.3 Example: Screening for Breast Cancer (cont).

- **Answer:** Consider the event $A = \text{"Woman has breast cancer"}$ and $B = \text{"Test positive"}$. We are interested in $P(A|B)$ where

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}.$$

Here we have

$$P(A) = 0.01 = 1 - P(A^c), \quad P(B|A) = 0.85, \quad P(B|A^c) = 0.10$$

so

$$P(A|B) = \frac{0.85 \times 0.01}{0.85 \times 0.01 + 0.10 \times 0.99} \approx 0.08.$$

- There is "only" 8% of chance the woman having breast cancer if the test is positive.

3.3 Example: Prosecutor's Fallacy

- A zealous prosecutor has collected an evidence for a crime and has an expert testify that the proba of finding this evidence if the accused were innocent is one-in-a-million and one if he is guilty. The prosecutor concludes that the probability of the accused being innocent is one-in-a-million. This is WRONG. Compute the probability that the accused is guilty assuming the population of the region where the crime has been committed is of 10 million people
- **Answer:** Define A = "The accused is guilty" and B = "Finding this evidence" then

$$P(A) = 10^{-7}, P(B|A) = 1 \text{ and } P(B|A^c) = 10^{-6}$$

so

$$\begin{aligned} P(A|B) &= \frac{P(B|A) P(A)}{P(B|A) P(A) + P(B|A^c) P(A^c)} \\ &= \frac{10^{-7}}{10^{-7} + 10^{-6} \times (1 - 10^{-7})} \approx 0.1. \end{aligned}$$

3.3 Example: Let's make a deal TV show

- Three doors A,B,C are closed. Behind one of them, there is a Ferrari, behind the others there is a goat. The player selects one door, say A for sake of simplicity. The TV presenter, who knows where the car is, tells him that the car is not behind door B and offers him the possibility to revise her/his choice (i.e. that is to select door C). Should the player revise his choice?
- **Answer:** Let A (resp. B , C) the event "the Ferrari is behind door A " (resp. B , C). Let E the event "the presenter tells the player that it is not behind door B ". We want to compute $P(A|E)$. We have

$$P(A) = P(B) = P(C) = \frac{1}{3}$$

and

$$P(E|A) = \frac{1}{2}, \quad P(E|B) = 0, \quad P(E|C) = 1.$$

Indeed, if the car is behind A then the presenter has the choice between B and C . If the car is behind C , the presenter can only pick door B as he/she cannot select A .

3.3 Example: Let's make a deal TV show (cont.)

- Hence we have

$$\begin{aligned}P(E) &= P(E|A)P(A) + P(E|B)P(B) + P(E|C)P(C) \\&= \frac{1}{2} \times \frac{1}{3} + \frac{1}{3} = \frac{1}{2}\end{aligned}$$

and

$$P(A|E) = \frac{P(E|A)P(A)}{P(E)} = \frac{\frac{1}{2} \times \frac{1}{3}}{\frac{1}{2}} = \frac{1}{3}$$

and thus

$$P(C|E) = 1 - P(A|E) = \frac{2}{3}.$$

- You should definitely revise your choice.

3.3 Example: A Prisoners' Dilemma

- Three prisoners, A, B and C, are in separate cells and sentenced to death. The governor has selected one of them at random to be pardoned. The warden knows which one is pardoned, but is not allowed to tell. Prisoner A begs the warden to let him know the identity of one of the others who is going to be executed. "If B is to be pardoned, give me C's name. If C is to be pardoned, give me B's name. And if I'm to be pardoned, flip a coin to decide whether to name B or C."
- (a). The warden tells A that B is to be executed. Prisoner A is pleased because he believes that his probability of surviving has gone up from $1/3$ to $1/2$, as it is now between him and C. Is it correct?
- (b). Prisoner A secretly tells C the news, who is also pleased, because he reasons that A still has a chance of $1/3$ to be the pardoned one, but his chance has gone up to $2/3$. Is it correct?

3.4 Independence

- We say that two events E, F are independent if

$$P(E \cap F) = P(E) P(F)$$

or equivalently

$$P(E|F) = P(E) \text{ and } P(F|E) = P(F).$$

- If E and F are independent then E and F^c are also independent as

$$\begin{aligned} P(E) &= P(E \cap F) + P(E \cap F^c) \\ &= P(E) P(F) + P(E \cap F^c) \end{aligned}$$

so

$$P(E \cap F^c) = P(E) (1 - P(F)) = P(E) P(F^c).$$

3.4 Example

- You toss two air dice. Let E the event that the sum of the dice is 8 and F the event that the first die equals 4. Are E and F independent?
- **Answer.** We have $E = (\{2, 6\}, \{3, 5\}, \{4, 4\}, \{5, 3\}, \{6, 2\})$ and $F = (\{4, 1\}, \{4, 2\}, \dots, \{4, 6\})$

$$P(E)P(F) = \frac{5}{36} \times \frac{1}{6} = \frac{5}{126}$$

but $E \cap F = \{4, 4\}$ and

$$P(E \cap F) = \frac{1}{36}.$$

So the events are not independent.

3.4 Independence

- We say that three events E, F, G are independent if

$$\begin{aligned}P(E \cap F \cap G) &= P(E) P(F) P(G), \\P(E \cap F) &= P(E) P(F), \quad P(E \cap G) = P(E) P(G), \\P(F \cap G) &= P(F) P(G).\end{aligned}$$

- This generalizes to any events $\{E_i\}_{i=1}^n$. For any $i_1, i_2, \dots, i_r$ where $i_j \neq i_k$ for $j, k \in \{1, \dots, r\}$ and any $r \in \{1, \dots, n\}$ we have

$$P\left(\cap \{E_{i_j}\}_{j=1, \dots, r}\right) = \prod_{j=1}^r P(E_{i_j}).$$

3.4 Example: Independence of Events

1. Suppose two events A and B are mutually exclusive. If $P(A) > 0$ and $P(B) > 0$, can A and B be independent?
2. Suppose

$$S = \{a, b, c, d\}$$

and each basic outcome is equally likely to occur. Let $A_1 = \{a, b\}$, $A_2 = \{b, c\}$, and $A_3 = \{a, c\}$. Are A_1 , A_2 and A_3 independent?

3. Suppose $S = \{a, b, c, d, e, f, g, h\}$ and each basic outcome is equally likely to occur. Let $A_1 = \{a, b, c, d\}$, $A_2 = \{a, b, c, d\}$, and $A_3 = \{a, e, f, g\}$. Are A_1 , A_2 and A_3 independent?

3.4 Example: Independent (Bernoulli) Trials

- An infinite sequence of independent trials is to be performed. With proba. p it is a success and with proba $1 - p$ this is a failure. What is the proba that 1) at least one success occurs in the first n trials 2) exactly k successes occur in the first n trials 3) all trials results in successes?
- **Answer:** 1) Let E_i event of a failure at trial i then proba of no success is

$$P(E_1 \cap E_2 \cap \cdots \cap E_n) = \prod_{i=1}^n P(E_i) = (1 - p)^n$$

so proba of at least one success is $1 - (1 - p)^n$.

2) Proba of exactly k success is

$$\binom{n}{k} p^k (1 - p)^{n-k}.$$

3) Proba of all success is

$$P(\cap_{i=1}^{\infty} E_i^c) = \lim_{n \rightarrow \infty} P(\cap_{i=1}^n E_i^c) = \lim_{n \rightarrow \infty} p^n = 0 \text{ for } p < 1.$$

3.4 Example: The problem of points

- An infinite sequence of independent trials is to be performed. With proba. p it is a success and with proba $1 - p$ this is a failure. What is the proba of getting at least n successes before getting m failures?
- **Answer:** For at least n successes to occur before m failures, we need at least n successes in the first $n + m - 1$ trials. Hence, we have

$$\begin{aligned} P(\text{at least } n \text{ successes}) &= \sum_{k=n}^{n+m-1} P(k \text{ successes exactly}) \\ &= \sum_{k=n}^{n+m-1} \binom{n+m-1}{k} p^k (1-p)^{m+n-1-k} \end{aligned}$$

3.4 Example: Dice... again

- Independent trials consisting of rolling a pair of fair dice are performed. What is the probability that an outcome of 5 appears before an outcome of 7 when the outcome is the sum of the dice?
- Answer:** Let $E_n =$
 $\{\text{no 5 and 7 appears in the first } n - 1 \text{ trials and 5 appears at } n.\}$
then

$$P(\cup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} P(E_n).$$

where $P(\text{get a 5}) = \frac{4}{36}$, $P(\text{get a 7}) = \frac{6}{36}$

$$\begin{aligned} P(E_n) &= P(\text{no 5, no 7 } n - 1 \text{ times}) P(\text{a 5 at } n) \\ &= \left(1 - \frac{10}{36}\right)^{n-1} \left(\frac{4}{36}\right) \end{aligned}$$

so finally

$$P(\cup_{n=1}^{\infty} E_n) = \frac{1}{9} \sum_{n=1}^{\infty} \left(\frac{13}{18}\right)^{n-1} = \frac{2}{5}.$$

3.4 Example: Game Strategies

- Two players have to cooperate to answer some questions. Both players are equally good. Given a question, they will answer correctly independently with probability p . Two strategies are proposed by the players.
- In Strategy 1, given a question, one selects randomly one of the two players by tossing a fair coin and his/her answer is chosen.
- In Strategy 2, the two players give independently their answers. If their answers agree, then this is the chosen answer. If they disagree, they toss a fair coin and pick randomly one of the answers.
- What is the best strategy?

3.4 Example: Game Strategies (cont.)

- Consider the event $R = \{\text{answer is correct}\}$,
 $P_1 = \{\text{Player 1 answers}\}$, $P_2 = P_1^c = \{\text{Player 2 answers}\}$. In the 1st strategy we have

$$\begin{aligned}P(R) &= P(R|P_1)P(P_1) + P(R|P_2)P(P_2) \\&= p \times \frac{1}{2} + p \times \frac{1}{2} = p.\end{aligned}$$

- In the 2nd strategy, we give an correct answer if
 $A_1 = \{\text{players agree and give correct answer}\}$, $A_2 = \{\text{player 1 right/player 2 wrong and toss decides player 1}\}$ or $A_3 = \{\text{player 1 wrong/player 2 right and toss decides player 2}\}$ occurs. Let A_4 denotes the event where the given answer is incorrect.

$$\begin{aligned}P(R) &= P(R \cap A_1) + P(R \cap A_2) + P(R \cap A_3) + P(R \cap A_4) \\&= p \times p + p \times (1 - p) \times \frac{1}{2} + (1 - p) \times p \times \frac{1}{2} = p.\end{aligned}$$

- Strategy 1 and Strategy 2 will give similar results.

3.5 Conditional Independence

- Given an event E and events $T_1, T_2, \dots, T_n$, the probability of E given $I_n = T_1 \cap T_2 \cap \dots \cap T_n$

$$P(E|I_n) = \frac{P(I_n|E)P(E)}{P(I_n|E)P(E) + P(I_n|E^c)P(E^c)}$$

- We say that the events $T_1, T_2, \dots, T_n$ are conditionally independent given E (respectively E^c) if for any positive integer $2 \leq r \leq n$, we have

$$P(T_{i_1} \cap T_{i_2} \cap \dots \cap T_{i_r} | E) = \prod_{j=1}^r P(T_{i_j} | E),$$

$$P(T_{i_1} \cap T_{i_2} \cap \dots \cap T_{i_r} | E^c) = \prod_{j=1}^r P(T_{i_j} | E^c).$$

3.5 Application to Spam Emails Detection

- **Example.** When you receive an email, your spam filter uses Bayes rule to decide whether it is spam or not. Basic spam filters check whether some pre-specified words appear in the email you have received; e.g. {diplomat,lottery,money,politics,president,sincerely,huge,work...}. That is to each email is associated n events T_i telling us whether the i th pre-specified word is in the email or not. Let $E = \{\text{email received is spam}\}$.
- Based on many (humanly) labelled training data, i.e. non-spam and spam emails, we estimate

$$P(E) = \text{proba. email is a spam,}$$

$$P(\text{word } i \text{ in} | E) = \text{proba. } i^{\text{th}} \text{ word in email given it is spam,}$$

$$P(\text{word } i \text{ in} | E^c) = \text{proba. } i^{\text{th}} \text{ word in email given it is not spam.}$$

- The spam filter makes the assumption that the events T_i are conditionally independent to compute $P(E | I_n)$ and decide whether it is spam or not.

3.5 Updating Information Sequentially

- **Example.** A patient came to a clinic. We want to assess whether he has a nasty disease. For this we need to run not one but a battery of $n > 1$ tests. Let introduce the event $A = \{\text{patient sick}\}$ and the events $T_i = \{\text{result test of } i \text{ is positive}\}$. Assuming for example that we have n tests and observed the events $T_1, T_2, \dots, T_n$ then we are interested in

$$P(A | I_n) = \frac{P(I_n | A) P(A)}{P(I_n | A) P(A) + P(I_n | A^c) P(A^c)}.$$

where $I_n = T_1 \cap T_2 \cap \dots \cap T_n$ is the result of all the n tests

- These tests might be very expensive so instead of running directly two tests, we might want to first compute $P(A | I_1) = P(A | T_1)$ after the 1st test which is given by

$$P(A | I_1) = \frac{P(T_1 | A) P(A)}{P(T_1 | A) P(A) + P(T_1 | A^c) P(A^c)},$$

then only running the 2nd test if $P(A | I_1)$ is not close enough to 0 or 1 so that we cannot decide whether the patient is healthy or sick.

3.5 Updating Information Sequentially

- Assume we run the 2nd test and thus observe T_2 , then we are now interested in $P(A|I_2)$ given by

$$P(A|I_2) = \frac{P(I_2|A)P(A)}{P(I_2|A)P(A) + P(I_2|A^c)P(A^c)}$$

If $P(A|I_2)$ is not close enough to 0 or 1 so that we cannot decide whether the patient is healthy or sick, then we run the 3rd test and compute $P(A|I_3)$ and so on.

- In many situations, we have conditional independence of the events

$$P(I_n|A) = P(T_1 \cap T_2 \cap \cdots \cap T_n|A) = \prod_{i=1}^n P(T_i|A),$$

$$P(I_n|A^c) = P(T_1 \cap T_2 \cap \cdots \cap T_n|A^c) = \prod_{i=1}^n P(T_i|A^c).$$

- In this case, we can compute $P(A|I_k)$ as a function of $P(A|I_{k-1})$.

3.5 Updating Information Sequentially

- We want to show that

$$P(A|I_k) = \frac{P(T_k|A)P(A|I_{k-1})}{P(T_k|A)P(A|I_{k-1}) + P(T_k|A^c)P(A^c|I_{k-1})}.$$

- The proof follows from

$$P(A|I_k) = \frac{P(I_k|A)P(A)}{P(I_k|A)P(A) + P(I_k|A^c)P(A^c)}$$

where $P(I_k|A) = P(T_k|A)P(I_{k-1}|A)$ and $P(I_k|A^c) = P(T_k|A^c)P(I_{k-1}|A^c)$. So we have

$$\begin{aligned} P(A|I_k) &= \frac{P(T_k|A)P(I_{k-1}|A)P(A)}{P(T_k|A)P(I_{k-1}|A)P(A) + P(T_k|A^c)P(I_{k-1}|A^c)P(A^c)} \\ &= \frac{P(T_k|A)P(I_{k-1}|A)P(A)/P(I_{k-1})}{P(T_k|A)P(I_{k-1}|A)P(A)/P(I_{k-1}) + P(T_k|A^c)P(I_{k-1}|A^c)P(A^c)/P(I_{k-1})} \end{aligned}$$

and the result follows.

3.5 Example: Conditional Independence and HIV Test

- Data from joint United Nation Programs on HIV/AIDS (2006): We have a test such that $P(T^+ | HIV^+) = 0.99$ (sensitivity) and $P(T^- | HIV^-) = 0.99$ (specificity). The prevalence of HIV in East Asia is 0.001. Given the first test is positive, event T_1^+ , what is the probability of being HIV positive? The policy for a positive HIV test is a follow-up confirmatory test. Given the 2nd test is positive, event T_2^+ , what is the proba of being HIV positive?
- **Answer.** We have

$$\begin{aligned} P(HIV^+ | T_1^+) &= \frac{P(T_1^+ | HIV^+)P(HIV^+)}{P(T_1^+ | HIV^+)P(HIV^+) + P(T_1^+ | HIV^-)P(HIV^-)} \\ &= \frac{0.99 \times 0.001}{0.99 \times 0.001 + 0.01 \times 0.999} = 0.0902 \end{aligned}$$

If the 2nd test is positive, we need to compute

$$P(HIV^+ | T_1^+ \cap T_2^+) = \frac{P(T_1^+ \cap T_2^+ | HIV^+)P(HIV^+)}{P(T_1^+ \cap T_2^+ | HIV^+)P(HIV^+) + P(T_1^+ \cap T_2^+ | HIV^-)P(HIV^-)}$$

3.5 Example: Conditional Independence and HIV Test

- Given the information we have been given, we can only compute this proba if we assume that

$$\begin{aligned}P(T_1^+ \cap T_2^+ | HIV^+) &= P(T_1^+ | HIV^+) P(T_2^+ | HIV^+), \\P(T_1^+ \cap T_2^+ | HIV^-) &= P(T_1^+ | HIV^-) P(T_2^+ | HIV^-)\end{aligned}$$

In this case we have

$$P(HIV^+ | T_1^+ \cap T_2^+) = \frac{0.99 \times 0.99 \times 0.001}{0.99 \times 0.99 \times 0.001 + 0.01 \times 0.01 \times 0.999} = 0.9075$$

- Alternatively, we can use the sequential updating rule

$$\begin{aligned}P(HIV^+ | T_1^+ \cap T_2^+) &= \frac{P(T_2^+ | HIV^+) P(HIV^+ | T_1^+)}{P(T_2^+ | HIV^+) P(HIV^+ | T_1^+) + P(T_2^+ | HIV^-) P(HIV^- | T_1^+)} \\&= \frac{0.99 \times 0.0902}{0.99 \times 0.0902 + 0.01 \times 0.9098} = 0.9075\end{aligned}$$