Bayes' formula
Independent events

February 27, 2019

Conditional probability

- Probability space (Ω, P)
- A, B are 2 events in Ω
- Suppose we know B happens
- If A happens then only samples in AB can happen

Conditional probablity

- B becomes a new sample space
- A has a new probability in new sample space

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

- called the *conditional probability of A given B*

Multiplication formula

From the definition of $P(A|B)$:

$$P(A_1A_2) = P(A_1)P(A_2|A_1)$$

Think of A_1A_2 as event with 2 steps, then probability equals probability of first step multiply with the conditional probability of second step given first step

Multistep

Sequence of events $A_1, A_2, \ldots, A_k$

$$\begin{aligned} P(A_1 A_2 \ldots A_k) =& P(A_1) \times P(A_2|A_1) \\ & \times P(A_3|A_1 A_2) \ldots \\ & \times P(A_k|A_1 \ldots A_{k-1}) \end{aligned}$$

Example

An ordinary deck of 52 playing cards is randomly divided into 4 piles of 13 cards each. Compute the probability that each pile has exactly 1 ace.

Solution

Define events $E_i, i = 1, 2, 3, 4,$:

$E_1 = \{$ace of spades is in any one of the piles$\}$

$E_2 = \{$ace of spades and ace of hearts are in different piles$\}$

$E_3 = \{$aces of spades, hearts, and diamonds are all in different piles$\}$

$E_4 = \{4$ aces are in different piles$\}$

Solution

- $P(E_1) = 1$
- $P(E_2|E_1) = (52 - 13)/(52 - 1) = 39/51$
- $P(E_3|E_1E_2) = (52 - 13 \times 2)/(52 - 2) = 26/50$
- $P(E_4|E_1E_2E_3) = (52 - 13 \times 3)/(52 - 3) = 13/49$

Solution

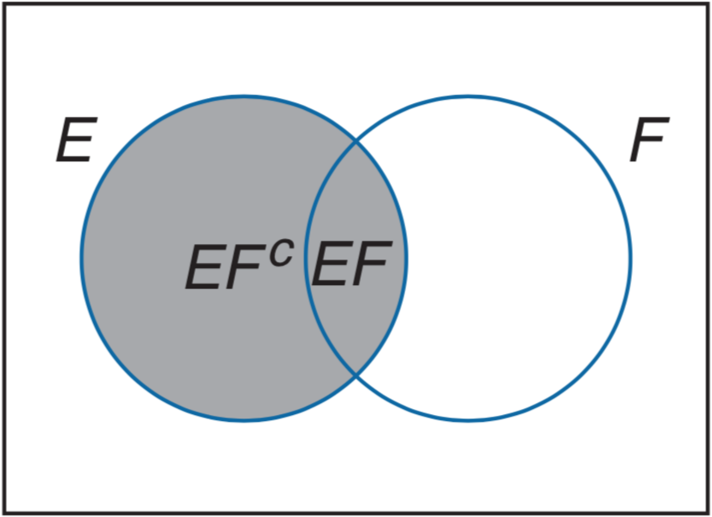
$$P(E_4) = P(E_1E_2E_3E_4)$$
$$= P(E_1)P(E_2|E_1)P(E_3|E_1E_2)P(E_4|E_1E_2E_3)$$
$$= 1 \cdot \frac{39}{51} \cdot \frac{26}{50} \cdot \frac{13}{49}$$

Total probability formula

Let E and F are 2 events. Then we may write

$$E = EF \cup EF^c.$$

EF and EF^c are mutually exclusive.


E
F
EF^c
EF

Total probability formula

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

How to use

- To find $P(E)$, need $P(E|F)$ and $P(E|F^c)$ and $P(F)$
- Assume that F happens, we can calculate $P(E|F)$
- Assume that F doesn't happen, calculate $P(E|F^c)$
- Usually $P(E|F)$ is easier to compute than $P(E)$ because we have more information about F

How to use

- To find $P(E)$, need $P(E|F)$ and $P(E|F^c)$ and $P(F)$
- Assume that F happens, we can calculate $P(E|F)$
- Assume that F doesn't happen, calculate $P(E|F^c)$
- Usually $P(E|F)$ is easier to compute than $P(E)$ because we have more information about F

How to use

- To find $P(E)$, need $P(E|F)$ and $P(E|F^c)$ and $P(F)$
- Assume that F happens, we can calculate $P(E|F)$
- Assume that F doesn't happen, calculate $P(E|F^c)$
- Usually $P(E|F)$ is easier to compute than $P(E)$ because we have more information about F

How to use

- To find $P(E)$, need $P(E|F)$ and $P(E|F^c)$ and $P(F)$
- Assume that F happens, we can calculate $P(E|F)$
- Assume that F doesn't happen, calculate $P(E|F^c)$
- Usually $P(E|F)$ is easier to compute than $P(E)$ because we have more information about F

Example

An insurance company thinks that there are some people who are accident prone and others who are not. An accident-prone person will have an accident at some time within 1 year with probability .4, whereas this probability is .2 for other person. Assume that 30 percent of the population is accident prone, what is the probability that a new policy holder will have an accident within a year of purchasing a policy?

Solution

- E ={new customer will have accident}
- F ={new customer is accident prone}
- $P(F) = 0.3$
- $P(E|F) = 0.4$, $P(E|F^c) = 0.2$

Solution

- E ={new customer will have accident}
- F ={new customer is accident prone}
- $P(F) = 0.3$
- $P(E|F) = 0.4$, $P(E|F^c) = 0.2$

Solution

- E ={new customer will have accident}
- F ={new customer is accident prone}
- $P(F) = 0.3$
- $P(E|F) = 0.4$, $P(E|F^c) = 0.2$

Solution

- E ={new customer will have accident}
- F ={new customer is accident prone}
- $P(F) = 0.3$
- $P(E|F) = 0.4$, $P(E|F^c) = 0.2$

Solution

By Bayes:

$$\begin{aligned} P(E) &= P(E|F)P(F) + P(E|F^c)P(F^c) \\ &= 0.4 \times 0.3 + 0.2 \times 0.7 \\ &= 0.26 \end{aligned}$$

Bayes formula

Change the order of condition

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

$$= \frac{P(F|E)P(E)}{P(E)P(F|E) + (1 - P(E))P(F|E^c)}$$

We can use $P(F|E)$ and $P(F|E^c)$ to calculate $P(E|F)$

Re-evaluate

For any new customer the company assumes that the probability that he is accident-prone is 0.3. There is one customer who had accident within 1 year. What is the probability that he is accident-prone given the new information?

Solution

A={the customer had accident} then $P(A) = P(E)$.

$$P(F|A) = \frac{P(FA)}{P(A)} = \frac{P(A|F)P(F)}{P(A)} = \frac{0.4 \times 0.3}{0.26} = 0.4615$$

Solution

A={the customer had accident} then $P(A) = P(E)$.

$$
\begin{aligned}
P(F|A) &= \frac{P(FA)}{P(A)} \\
&= \frac{P(A|F)P(F)}{P(A)} \\
&= \frac{0.4 \times 0.3}{0.26} = 0.4615
\end{aligned}
$$

Example

In answering a question on a multiple choice test, let p be the probability that the student knows the answer and $1 - p$ the probability that she guesses. Assume that a student who guesses at the answer will be correct with probability $1/m$, where m is the number of multiple-choice alternatives. What is the conditional probability that a student knew the answer to a question given that she answered it correctly?

Solution

- A={Student knows the answer}
- B={Student answered correctly}
- $P(A) = p$, $P(B|A^c) = 1/m$, $P(A|B) =?$

Solution

- A={Student knows the answer}
- B={Student answered correctly}
- $P(A) = p$, $P(B|A^c) = 1/m$, $P(A|B) = ?$

Solution

- A={Student knows the answer}
- B={Student answered correctly}
- $P(A) = p$, $P(B|A^c) = 1/m$, $P(A|B) = ?$

Solution

$$
\begin{aligned}
P(A|B) &= \frac{P(B|A)P(A)}{P(B)} \\
&= \frac{1 \times p}{P(B|A)P(A) + P(B|A^c)P(A^c)} \\
&= \frac{p}{1 \times p + (1/m) \times (1-p)} \\
&= \frac{mp}{1 + (m-1)p}
\end{aligned}
$$

Practice

A laboratory blood test is 99 percent effective in detecting a certain disease when it is present, but it also gives a ”false positive” result for 1 percent of the healthy persons tested. (If a healthy person is tested, then, with probability .01, the test result will say that he has the disease.) If .5 percent of the population actually has the disease, what is the probability a person has the disease given that his test result is positive?

Partition of sample space

$F_1, \ldots, F_n$ is a partition of Ω: they are mutually exclusive and their union is Ω. Then for any event E

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

General Bayes' Formula

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

We can use this to update the probability of F_i given the new evident E.

General Bayes' Formula

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

We can use this to update the probability of F_i given the new evident E.

Example

A plane is missing and it was equally likely to have gone down in any of three possible regions. Let $1 - \alpha_i$ denote the probability the plane will be found upon a search of the i-th region when the plane is, in fact, in that region, $i = 1, 2, 3$. What is the conditional probability that the plane is in the i-th region, given that a search of region 1 is unsuccessful, $i = 1, 2, 3$?

Solution

- F_i={the plane is in region i}
- E={search of region 1 was unsuccessful}
- $P(F_i) = 1/3$. Need $P(F_i|E) =?$

Solution

- F_i={the plane is in region i}
- E={search of region 1 was unsuccessful}
- $P(F_i) = 1/3$. Need $P(F_i|E) = ?$

Solution

- F_i={the plane is in region i}
- E={search of region 1 was unsuccessful}
- $P(F_i) = 1/3$. Need $P(F_i|E) =?$

Solution

$$
\begin{aligned}
P(E) = P(E|F_1)P(F_1) + P(E|F_2)P(F_2) \\
+ P(E|F_3)P(F_3) \\
= \alpha_1 \times \frac{1}{3} + 1 \times \frac{1}{3} + 1 \times \frac{1}{3}
\end{aligned}
$$

Solution

$$
\begin{aligned}
P(E) = P(E|F_1)P(F_1) &+ P(E|F_2)P(F_2) \\
&+ P(E|F_3)P(F_3) \\
= \alpha_1 \times \frac{1}{3} + 1 \times \frac{1}{3} + 1 \times \frac{1}{3}
\end{aligned}
$$

Solution

$$P(F_1|E) = \frac{P(E|F_1)P(F_1)}{P(E)} = \frac{\alpha_1}{\alpha_1 + 2}$$

$$P(F_2|E) = \frac{P(E|F_2)P(F_2)}{P(E)} = \frac{1}{\alpha_1 + 2}$$

Solution

$$P(F_1|E) = \frac{P(E|F_1)P(F_1)}{P(E)} = \frac{\alpha_1}{\alpha_1 + 2}$$

$$P(F_2|E) = \frac{P(E|F_2)P(F_2)}{P(E)} = \frac{1}{\alpha_1 + 2}$$

Independent events

Idea

- Usually $P(E|F) \neq P(E)$.
- If $P(E|F) = P(E)$, F has no effect on E
- E and F have no relation

Idea

- Usually $P(E|F) \neq P(E)$.
- If $P(E|F) = P(E)$, F has no effect on E
- E and F have no relation

Idea

- Usually $P(E|F) \neq P(E)$.
- If $P(E|F) = P(E)$, F has no effect on E
- E and F have no relation

Independent events

If $P(E|F) = P(E)$ then

$$P(EF) = P(E)P(F).$$

In this case we say E and F are independent events.

Independent events

If $P(E|F) = P(E)$ then

$$P(EF) = P(E)P(F).$$

In this case we say E and F are independent events.

- Mutually exclusive $\neq$ independent
- Ex: E and E^c are mutually exclusive but not independent.
- To understand independency, have to think of E and F as real world events, not subsets.

- Mutually exclusive $\neq$ independent
- Ex: E and E^c are mutually exclusive but not independent.
- To understand independency, have to think of E and F as real world events, not subsets.

- Mutually exclusive $\neq$ independent
- Ex: E and E^c are mutually exclusive but not independent.
- To understand independency, have to think of E and F as real world events, not subsets.

Example

A card is selected at random from an ordinary deck of 52 playing cards. Let A be the event that the selected card is an ace and H be the event that it is a heart, then are A and H independent?

- $P(A) = 4/52$
- $P(H) = 13/52$
- $P(AH) = 1/52 = P(A)P(H)$
- A and H are independent (but not mutually exclusive)

- $P(A) = 4/52$
- $P(H) = 13/52$
- $P(AH) = 1/52 = P(A)P(H)$
- A and H are independent (but not mutually exclusive)

- $P(A) = 4/52$
- $P(H) = 13/52$
- $P(AH) = 1/52 = P(A)P(H)$
- A and H are independent (but not mutually exclusive)

- $P(A) = 4/52$
- $P(H) = 13/52$
- $P(AH) = 1/52 = P(A)P(H)$
- A and H are independent (but not mutually exclusive)

Complement

If E is independent of F then it is independent of F^c.

3 events

Suppose E is independent of F and also is independent of G. Are E and FG independent?

Example

Roll 2 dice.

- E={sum is 7}
- F={1st dice is 4}
- G={2nd dice is 3}
- E independent of F and of G but $P(E|FG) = 1$

Example

Roll 2 dice.

- E={sum is 7}
- F={1st dice is 4}
- G={2nd dice is 3}
- E independent of F and of G but $P(E|FG) = 1$

3 independent events

E, F, G are independent if

- $P(EFG) = P(E)P(F)P(G)$
- $P(EF) = P(E)P(F)$
- $P(FG) = P(F)P(G)$
- $P(EG) = P(E)P(G)$

If E, F, G are independent then E is independent of any event formed by F and G: $F \cup G$, FG, ...

Independent experiments

- Toss n coins: all coins are independent
- Students do exam in separate rooms: results are independent.

Independent experiments

- Toss n coins: all coins are independent
- Students do exam in separate rooms: results are independent.

Example

k coupons is collected, each of which is independently a type j coupon with probability $p_j, j = 1, .., n,$

$$\sum_{j=1}^{n} p_j = 1.$$

Find the probability that the set contains a type j coupon given that it contains a type i, $i \neq j$.

Solution

A_j={ set contains at least one type j}

A_j^c ={no type j coupon}

$$P(A_j^c) = (1 - p_j)^k$$

b/c each of k coupon is independent

$$P(A_j) = 1 - P(A_j^c) = 1 - (1 - p_j)^k$$

$$
\begin{aligned}
P(A_j A_i) &= 1 - P(A_j^c \cup A_i^c) \\
&= 1 - P(AJ^c) - P(A_i^c) + P(A_j^c A_i^c) \\
&= 1 - (1 - p_j)^k - (1 - p_i)^k \\
&\quad + (1 - p_j - p_i)^k
\end{aligned}
$$

$$
P(A_j | A_i) =
$$

$$
= \frac{1 - (1 - p_j)^k - (1 - p_i)^k + (1 - p_j - p}{1 - (1 - p_i)^k}
$$

Practice

Each of 2 balls is painted black or gold and then placed in an urn.Suppose that each ball is colored black with probability $1/2$, and that these events are independent.

(a) Suppose that there is at least one of the balls is gold. Compute the conditional probability that both balls are gold.

(b) Suppose you take out one ball and it is gold. What is the probability that both balls are gold now?

Practice

A bin contains 3 types of flashlights: 20% are type 1, 30% are type 2, and 50 % are type 3. The probability that a type 1 flashlight will last over 100 hours of use is .7, with the corresponding probabilities for type 2 and type 3 being .4 and .3.

(a) Find the probability that a randomly chosen flashlight will last over 100 hours?

(b) Given that a flashlight lasted over 100 hours, what is the conditional probability that it was a type j flashlight, $j = 1, 2, 3$?

Homework 3

Chapter 3: 29, 30, 34, 36, 39