Chapter 3: Conditional Probability and Independence

Xiugang Wu

University of Delaware

Spring, 2020

Example

Roll two dice:

sum of 2 dice	2	3	4	5	6	7	8	9	10	11	12
probability	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$
prob. given 1st is 6	0	0	0	0	0	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Outline

- Conditional Probability
- Bayes' Formula
- Independent Events

Outline

- Conditional Probability
- Bayes' Formula
- Independent Events

Conditional Probability

The conditional probability that E occurs given that F has occurred is:

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

Interpretation: If event F has occurred, then in order for E to occur it is necessary that the actual occurrence is a point in both E and F, i.e. in EF. Now as F has occurred, F becomes our new (reduced) sample space; hence $P(E|F) = \frac{P(EF)}{P(F)}$.

If all outcomes are equally likely, then

$$P(E|F) = \frac{P(EF)}{P(F)} = \frac{\text{number of outcomes in } EF/\text{total number of outcomes}}{\text{number of outcomes in } F/\text{total number of outcomes}}$$
$$= \frac{\text{number of outcomes in } EF}{\text{number of outcomes in } F}$$

Example

Example: Two dice are rolled. If the first dice is 6, what is the probability that the sum of the dice is 7?

Solution: Let $F = \{\text{the first dice is } 6\}$ and $E = \{\text{the sum of the dice is } 7\}$. Then we have

$$P(E|F) = \frac{P(EF)}{P(F)} = \frac{P(\{(6,1)\})}{P(F)} = \frac{1/36}{6/36} = \frac{1}{6}.$$

Example: A student is taking a one-hour-limit-exam. Suppose the probability that the student will finish the exam in less than x hour is $x/2$ for $x \in [0,1]$. Given that the student is still working after 0.75 hour, what is the probability that the full hour is needed?

Solution: Let $F = \{\text{more than } 0.75 \text{ hour is needed}\}$ and $E = \{\text{the full hour is needed}\}$. Then we have $P(E) = (1 - x/2)|_{x=1} = 0.5$ and $P(F) = (1 - x/2)|_{x=0.75} = 0.625$, and hence

$$P(E|F) = \frac{P(EF)}{P(F)} = \frac{P(E)}{P(F)} = \frac{0.5}{0.625} = 0.8.$$

Generalization

Conditional on multiple events:

$$P(E|F_1 \cdots F_n) = \frac{P(EF_1 \cdots F_n)}{P(F_1 \cdots F_n)}.$$

Multiplication rule:

$$P(E_1 E_2 \cdots E_n) = P(E_1)P(E_2|E_1) \cdots P(E_n|E_1 \cdots E_{n-1}).$$

Example

Example: 52 playing cards are randomly divided into 4 piles of 13 cards each. What is the probability that each pile has exactly one ace?

Solution 1:

$$\frac{4!\binom{48}{12,12,12,12}}{\binom{52}{13,13,13,13}} \approx 0.105.$$

Example

Example: 52 playing cards are randomly divided into 4 piles of 13 cards each. What is the probability that each pile has exactly one ace?

Solution 2: Let the four aces be Ace $1, 2, 3, 4$ and let $E_i = \{\text{pile } i \text{ has Ace } i\}$. Then

$$
\begin{aligned}
P(\{\text{each pile has one ace}\}) &= 4! P(E_1 E_2 E_3 E_4) \\
&= 4! P(E_1) P(E_2 | E_1) P(E_3 | E_1 E_2) P(E_4 | E_1 E_2 E_3).
\end{aligned}
$$

To calculate $P(E_1)$ one can think of randomly throwing Ace 1 into 52 holes

$$
\underbrace{\sqcup \sqcup \cdots \sqcup}_{\text{pile 1}}\underbrace{\sqcup \sqcup \cdots \sqcup}_{\text{pile 2}}\underbrace{\sqcup \sqcup \cdots \sqcup}_{\text{pile 3}}\underbrace{\sqcup \sqcup \cdots \sqcup}_{\text{pile 4}}
$$

in which process E_1 happens if Ace 1 lands on the first 13 holes; therefore $P(E_1) = \frac{13}{52}$. Now given E_1, one of the holes in pile 1 has been occupied, and the probability Ace 2 lands on pile 2 is $P(E_2 | E_1) = \frac{13}{51}$. Similarly, we have $P(E_3 | E_1 E_2) = \frac{13}{50}$ and $P(E_4 | E_1 E_2 E_3) = \frac{13}{49}$.

Outline

- Conditional Probability
- Bayes' Formula
- Independent Events

Bayes' Formula

The Law of Total Probability:

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

Bayes' Formula:

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

Example

Example: Suppose there are two types of people: accident-prone people, who have accident in one year with $p_1 = 0.4$; and non-accident-prone people, who have accident in one year with $p_2 = 0.2$. Assume 30% of the population is accident prone. What is the probability that a new policyholder will have an accident within a year of purchasing a policy?

Solution: Let $E = \{\text{new policyholder will have an accident}\}$, $F = \{\text{new policy-holder is accident-prone}\}$ and $F^c = \{\text{new policyholder is non-accident-prone}\}$. Then we have

$$\begin{aligned} P(E) &= P(E|F)P(F) + P(E|F^c)P(F^c) \\ &= 0.4 \times 0.3 + 0.2 \times (1 - 0.3) = 0.12 + 0.14 = 0.26 \end{aligned}$$

Suppose that the new policyholder has an accident in one year. The probability that he or she is accident-prone is

$$P(F|E) = \frac{P(EF)}{P(E)} = \frac{P(F)P(E|F)}{P(E)} = \frac{0.3 \times 0.4}{0.26} = \frac{6}{13}$$

Example

Example: In answering a question on a multiple-choice test with m choices, with probability p the student knows the answer, and with probability $1-p$ the student doesn't know the answer and randomly guesses the answer (with correct probability $1/m$). If the student answers right, what is the probability that the student really knows the answer?

Solution: Let $E = \{\text{he/she really knows the answer}\}$ and $F = \{\text{he/she answers right}\}$. Then

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

Generalization

Generally, suppose that $F_1, F_2, \cdots, F_n$ are mutually exclusive events such that $\cup_{i=1}^n F_i = S$. Then $E = \cup_{i=1}^n EF_i$ and

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

Bayes' Formula:

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

Example

Example: Three cards are identical in form. 1st card: both sides are red; 2nd card: both sides are black; 3rd card: one side is red and the other side is black. Randomly choose one card and throw it to the ground. If the upper side is red, what is the probability that the other side is black?

Solution: Let $E = \{\text{the upper side is red}\}$ and $F_i = \{i\text{th card is selected}\}$. Then the desired probability is

$$\begin{aligned} P(F_3|E) &= \frac{P(F_3)P(E|F_3)}{P(F_1)P(E|F_1) + P(F_2)P(E|F_2) + P(F_3)P(E|F_3)} \\ &= \frac{\frac{1}{3} \cdot \frac{1}{2}}{\frac{1}{3} \cdot 1 + \frac{1}{3} \cdot 0 + \frac{1}{3} \cdot \frac{1}{2}} \\ &= \frac{1}{3} \end{aligned}$$

Outline

- Conditional Probability
- Bayes' Formula
- Independent Events

Independent Events

We have seen that $P(E|F) \neq P(E)$ in general. However, sometimes we indeed have $P(E|F) = P(E)$; in this case we say E, F are independent.

Note that independence is mutual, because

$$
\begin{aligned}
\text{independence between } E \text{ and } F &\Leftrightarrow P(E|F) = P(E) \\
&\Leftrightarrow P(EF) = P(E)P(F) \\
&\Leftrightarrow P(F|E) = P(F)
\end{aligned}
$$

Proposition: If E and F are independent, then E and F^c are also independent.

Proof: Note that

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

Example

Example: Roll two dice. Let

$$E = \{\text{the sum of the two dice is 7}\}$$

and

$$F = \{\text{the first dice is 6}\}.$$

Then $P(E) = \frac{1}{6}$ and $P(E|F) = \frac{1}{6}$. Since $P(E) = P(E|F)$, the events E and F are independent.

Now consider $G = \{\text{the sum of the two dice is 5}\}$. Since $P(G) = \frac{4}{36}$ and $P(G|F) = 0$, the events G and F are not independent.

Example: Randomly select a card from 52 playing cards. Let

$$E = \{\text{selected card is an ace}\}$$

and

$$F = \{\text{selected card is a spade}\}.$$

Then $P(E) = \frac{4}{52}$, $P(F) = \frac{13}{52}$ and $P(EF) = \frac{1}{52}$. Since $P(EF) = P(E)P(F)$, the events E and F are independent.

Independence Among Three Events

Things can become tricky if there are three events E, F, G. Suppose E is independent of F, and E is independent of G. Is E necessarily independent of FG?

Answer: NO!

Example: Roll two dice. Let $E = \{\text{the sum is } 7\}$, $F = \{\text{the first dice is } 4\}$, and $G = \{\text{the second dice is } 3\}$. Show that E is independent of F and that E is independent of G. Now, is E independent of FG? No, because $P(E|FG) = 1$ which is not equal to $P(E)$.

Independence Among Three Events

Three events E, F, G are said to be independent if

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

This definition ensures that $P(E|FG) = P(E)$. Also we can obtain that $P(E|F \cup G) = P(E)$. Moreover, we can show that E is independent of any combination of events F and G.

Generally, $E_1, E_2, \ldots, E_n$ are independent if

$$P(E_{i_1} E_{i_2} \cdots E_{i_r}) = P(E_{i_1})P(E_{i_2}) \cdots P(E_{i_r})$$

for any subset $\{i_1, i_2, \ldots, i_r\} \subseteq \{1, 2, \ldots, n\}$.

Example

Example: An infinite sequence of independent trials is to be performed. Each trial results in a success with probability p. What is the probability that

(a) at least one success occurs in the first n trials;

(b) exactly k successes occur in the first n trials;

(c) all trials result in successes ?

Solution: (a) It is easier to calculate the probability that no success occurs in the first n trials. Let E_i = {failure in the ith trial}. Then $E_1 E_2 \cdots E_n$ = {no success in the first n trials}, and

$$P(E_1 E_2 \cdots E_n) = P(E_1)P(E_2) \cdots P(E_n) = (1-p)^n.$$

Therefore,

$$P(\{\text{at least one success}\}) = 1 - P(E_1 E_2 \cdots E_n) = 1 - (1-p)^n.$$

Example

Example: An infinite sequence of independent trials is to be performed. Each trial results in a success with probability p. What is the probability that

(a) at least one success occurs in the first n trials;

(b) exactly k successes occur in the first n trials;

(c) all trials result in successes ?

(b) Consider any particular sequence of the first n outcomes containing k successes and $n-1$ failures. According to independence, each one of these sequences will occur with probability $p^k(1-p)^{n-k}$. As there are $\binom{n}{k}$ such sequences, the desired probability is

$$P(\{\text{exactly } k \text{ successes}\}) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Example

Example: An infinite sequence of independent trials is to be performed. Each trial results in a success with probability p. What is the probability that

(a) at least one success occurs in the first n trials;

(b) exactly k successes occur in the first n trials;

(c) all trials result in successes ?

(c) The probability of the first n trials all resulting in success is given by $P(E_1^c E_2^c \cdots E_n^c) = p^n$. Therefore,

$$P(\{\text{all trials result in successes}\}) = \lim_{n \to \infty} p^n,$$

which equals 0 if $p < 1$ and 1 if $p = 1$.