Business Statistics with Python – Module BUSN969

Shaomin Wu

University of Kent

Art. No xxxxx
ISBN xxx–xx–xxxx–xx–x
Edition 0.0

Chapter 9

Probability and Bayes theorem

This chapter introduces basic probability concepts.

9.1 Probability

Definition 9.1 (Probability) Probability is the ratio of the number of favourable results to the total number of results.

Property 9.1 $0 \leq P(A) \leq 1$

9.1.1 Sample space

Definition 9.2 (sample space) The sample space of an experiment is the set of all possible results of that experiment.

The sample space is denoted by Ω.

Property 9.2 $P(\Omega) = 1$

9.1.2 Null event

Example 9.1 (Null event) Figure 9.1 shows a playing card that does not exist in a non-defective pack of playing cards as it is both the queen of diamonds and clubs. Hence, the event that there is such a card is a **null event**.

The null event is usually denoted by Φ.

Property 9.3 The probability of the null event is zero, that is, $P(\Phi) = 0$.

In probability, the Venn diagram is normally used to illustrate the relationship between events. A Venn diagram shows all possible logical relations between a finite number of different events.

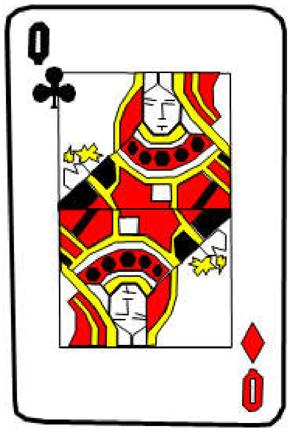

Figure 9.1: Null Event

9.1.3 Complementary events

Definition 9.3 (Complementary event) The complement of event A is defined to be the event consisting of all sample points that are not in A.

The complementary event of A is usually denoted by $\bar{A}$.

For example, the grey area in Figure 9.2 is the complementary event of the event A, which is the area in blue.

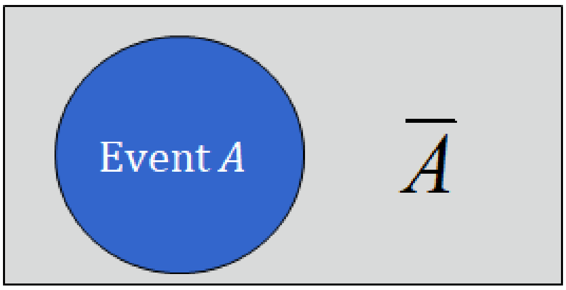

Figure 9.2: Complementary Event

Property 9.4 The probability of the complementary event of A is : $P(\bar{A}) = 1 - P(A)$.

9.1.4 Intersection of two events

Definition 9.4 The intersection of events A and B is the set of all sample points that are in both A and B.

The intersection of events A and B is denoted by $A \cap B$, or $A\&B$.

For example, the red area in Figure 9.3 is the intersection of event A and event B. Note: the red area is covered by both event A and event B.

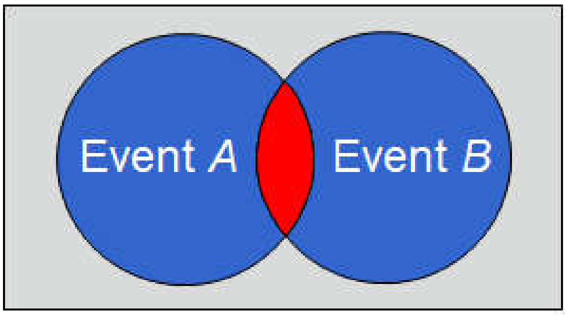

Figure 9.3: Intersection of event A and B

9.1.5 Union of two events

Definition 9.5 (Union of two events) The union of events A and B is the event containing all sample points that are in A or B or both.

The union of events A and B is denoted by $A \cup B$, A or B.

For example, the blue area (including the area in dark blue) in Figure 9.4 is the union of event A and event B.

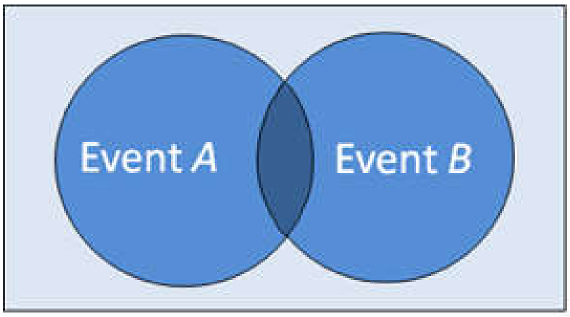

Figure 9.4: Union of Events A and B

The probability of $A \cup B$, $P(A \cup B)$, can be regarded as the area covered by both event A and event B.

Theorem 9.1 (Addition law) The probability of $A \cup B$ is given by

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{9.1}$$

The addition law in Eq. (9.1) *can be understood by this way: From Figure 9.4, the blue area, including the dark blue area, equals to the area covered by event A plus the area covered by event B, minus the area covered by both event A and event B.*

9.1.6 Mutually exclusive events

Definition 9.6 (Mutually exclusive events) Two events are mutually exclusive if, when one event occurs, the other cannot occur.

For example, the two events A and B in Figure 9.5 are mutually exclusive events as there is no area covered by both events.

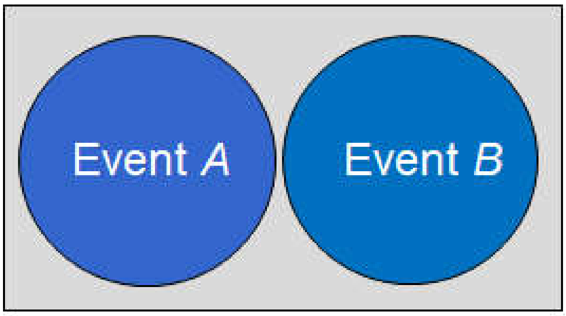

Figure 9.5: Events A and B are mutually exclusive events.

Theorem 9.2 If events A and B are mutually exclusive, $P(A \cap B) = 0$.

From both Eq. (9.1) and Eq. (9.2), one obtains

$$P(A \cup B) = P(A) + P(B) \tag{9.2}$$

9.1.7 Conditional probability

Definition 9.7 (Conditional probability) The probability of an event given that another event has occurred is called a conditional probability.

The conditional probability of A given B is denoted by $P(A|B)$.
The conditional probability can be obtained by

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{9.3}$$

or

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \tag{9.4}$$

From Eq. (9.3) and Eq. (9.4), we can obtain the multiplication law.

Theorem 9.3 (Multiplication Law) The probability of $A \cup B$ is given by

$$P(A \cap B) = P(A)P(B|A) = P(B)P(A|B) \tag{9.5}$$

9.1.8 Independent events

Definition 9.8 (Independent events) If the probability of event A is not changed by the existence of event B, we would say that events A and B are independent.

That is, events A and B are independent if

$$P(A|B) = P(A), \tag{9.6}$$

or

$$P(B|A) = P(B). \tag{9.7}$$

The multiplication law for Independent events becomes

$$P(B \cap A) = P(A)P(B). \tag{9.8}$$

Difference between mutually exclusive events and independent events.

- *Events A and B being mutually exclusive events suggests that if one of the events occurs, the other will not. This relationship can be illustrated by a Venn diagram or their probabilities:*
  - *There is no overlapping area between events A and B, see Fig 9.5;*
  - $P(A \cap B) = 0$.
- *Events A and B being independent events suggests that the occurrence of one event has no effect on that of the other one. This relationship can be illustrated by their probabilities, but not by a Venn diagram:*
  - $P(A \cap B) = P(A)P(B)$, *i.e.*, $P(A|B) = P(A)$ *and* $P(B|A) = P(B)$.

Example 9.2 (How much can Mick earn from his investment?) Mick has invested in two stocks, Smith Oil and David Mining. He estimated that the possible outcomes of these investments three months from now were shown in Table 9.1 and their probabilities are shown in Table 9.2. Mick has the following questions.

Q1: What is the probability that Smith Oil and David Mining will be profitable?

Q2: What is the probability that Smith Oil or David Mining will be profitable?

Q3: Is the profitability of Smith Oil independent of that of David Mining?

Q4: What is the probability that David Mining is profitable given Smith Oil being profitable?

Solution: Denote S={Smith Oil will be profitable} and D={David Mining will be profitable}. Then from Table 9.2,

$$S = \{(10,8),(10,-2),(5,8),(5,-2)\},$$

$$D = \{(10,8),(5,8),(0,8),(-20,8)\},$$

$$S \cup D = \{(10,8),(10,-2),(5,8),(5,-2),(0,8),(-20,8)\},$$

Table 9.1: Gain or loss (in £1,000).

Smith Oil	David Mining
10	8
5	-2
0	
-20	

Table 9.2: Gain or loss with probability

(10, 8)	18000	Gain	0.20
(10, -2)	8000	Gain	0.08
(5, 8)	13000	Gain	0.16
(5, -2)	3000	Gain	0.26
(0, 8)	8000	Gain	0.10
(0, -2)	2000	Loss	0.12
(-20, 8)	12000	Loss	0.02
(-20, -2)	22000	Loss	0.06

and

$$S \cap D = \{(10,8),(5,8)\}.$$

Hence,

$$\begin{aligned} P(S) &= P(10,8) + P(10,-2) + P(5,8) + P(5,-2) \\ &= 0.20 + 0.08 + 0.16 + 0.26 \\ &= 0.70, \end{aligned} \tag{9.9}$$

$$\begin{aligned} P(D) &= P(10,8) + P(5,8) + P(0,8) + P(-20,8) \\ &= 0.20 + 0.16 + 0.10 + 0.02 \\ &= 0.48, \end{aligned} \tag{9.10}$$

- To answer Q1, we need to find $P(S \cap D)$, which can be obtained by calculating it from Table 9.2

$$P(S \cap D) = P(10,8) + P(5,8) = 0.36. \tag{9.11}$$

- To answer Q2, we need to find $P(S \cup D)$, which can be obtained by calculating it from Table 9.2

$$\begin{aligned} P(S \cup D) &= P(10,8) + P(10,-2) + P(5,8) + P(5,-2) + P(0,8) + (-20,8) \\ &= 0.2 + 0.08 + 0.16 + 0.26 + 0.10 + 0.02 \\ &= 0.82, \end{aligned} \tag{9.12}$$

  or using Eq. (9.1),

$$\begin{aligned} P(S \cup D) &= P(S) + P(D) - P(S \cap D) \\ &= 0.70 + 0.48 - 0.36 \\ &= 0.82, \end{aligned} \tag{9.13}$$

- To answer Q3, based on Eq. (9.8), we need to prove $P(S \cap D) = P(S)P(D)$. Since $P(S \cap D) = 0.36$, $P(S) = 0.70$, and $P(D) = 0.48$, $0.36 \neq 0.70 \times 0.48$, that is, $P(S \cap D) \neq P(S)P(D)$.. We can therefore conclude that the profitability of Smith Oil is not independent of that of David Mining.

- Q4 asks you to provide $P(D|S)$. With Eq. (9.6) (or (9.7)), we have

$$\begin{aligned} P(D|S) &= \frac{P(S \cap D)}{P(S)} \\ &= 0.36/0.70 \\ &= 0.5143. \end{aligned} \tag{9.14}$$

9.1.9 Law of total probability

Theorem 9.4 (Law of total probability) Assume

(a) $(A_i \cap B) \cap (A_j \cap B) = \Phi$, for $i \neq j$, and

(b) $B = (A_1 \cap B) \cup (A_2 \cap B) \cup \cdots \cup (A_n \cap B)$,

then,

$$\begin{aligned} P(B) &= \sum_{i=1}^{n} P(A_i \cap B) \\ &= \sum_{i=1}^{n} P(A_i)P(B|A_i) \end{aligned} \tag{9.15}$$

where $\sum$ is a summation sign. For example, $\sum_{i=1}^{4} x_i = x_1 + x_2 + x_3 + x_4$ and $\sum_{i=1}^{n} x_i = x_1 + x_2 + \cdots + x_n$. Suppose there is an event B, as shown in Fig 9.6.

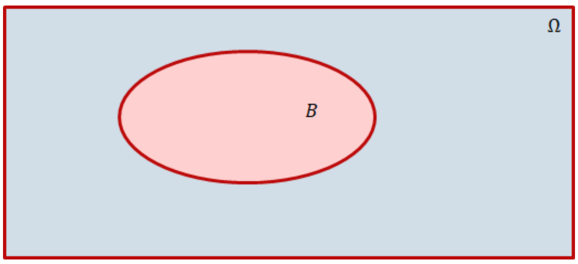

Figure 9.6: Event B in the sample space Ω

We cut B into n slices, as shown in Fig. 9.7. Then $(A_i \cap B) \cap (A_j \cap B) = \Phi$, for $i \neq j$. The area of event B is the sum of the areas of events $A_1 \cap B$, $A_2 \cap B$, …, and $A_n \cap B$. That is, $B = (A_1 \cap B) \cup (A_2 \cap B) \cup \cdots \cup (A_n \cap B)$.

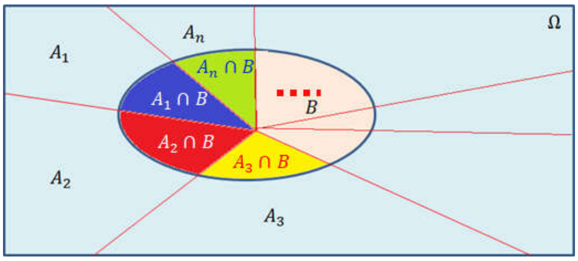

Figure 9.7: Event B is split by n events $A_1, A_2, ..., A_n$, $(A_i \cap B) \cap (A_j \cap B) = \Phi$, for $i \neq j$.

Then

$$\begin{aligned} P(B) &= P((A_1 \cap B) \cup (A_2 \cap B) \cup \cdots \cup (A_n \cap B)) \\ &= P(A_1 \cap B) + P(A_2 \cap B) + \cdots + P(A_n \cap B)) \\ &= \sum_{i=1}^{n} P(A_i \cap B) \qquad (9.16) \\ &= \sum_{i=1}^{n} P(A_i)P(B|A_i). \end{aligned}$$

With Eq. (9.6), $P(A_i \cap B) = P(A_i)P(B|A_i)$. Then we have

$$\begin{aligned} P(B) &= \sum_{i=1}^{n} P(A_i \cap B) \\ &= \sum_{i=1}^{n} P(A_i)P(B|A_i). \end{aligned} \qquad (9.17)$$

9.1.10 Bayes' Theorem

Note

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(B|A_i)P(A_i)}{P(B)}. \qquad (9.18)$$

Substituting $P(B)$ in Eq. (9.17) into Eq. (9.18), we can obtain the following theorem, or Bayes's theorem.

Theorem 9.5 **(Bayes's theorem)** The probability that event A_i will occur given that event B has occurred is obtained by

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)}. \qquad (9.19)$$

In Bayes theorm (9.19), $P(A_i)$, $P(B|A_i)$, and $P(A_i|B)$ are referred to as the prior distribution, the likelihood, and the posterior distribution, respectively.