Law of Total Probability **Given a sequence of mutually exclusive events** $S_1, S_2, \cdots, S_n$. **If event** $A \subset \cup_{i=1}^{n} S_i$ **and** $\mathbb{P}(S_i) > 0$, **then**

$$\mathbb{P}(A) = \mathbb{P}(S_1)\mathbb{P}(A|S_1) + \cdots + \mathbb{P}(S_n)\mathbb{P}(A|S_n)$$

Proof: **Since** $(A \cap S_i) \cap (A \cap S_j) = \phi$ **for any** $i \neq j$, **this means that the sequence** $\{(A \cap S_i), i = 1, \cdots, n\}$ **are mutually exclusive, according to the second condition of the definition of probability, we have**

$$\mathbb{P}(A) = \mathbb{P}(A \cap S_1) + \cdots + \mathbb{P}(A \cap S_n).$$

Since

$$\mathbb{P}(A \cap S_i) = \mathbb{P}(S_i)\mathbb{P}(A|S_i) \ \textbf{for} \ i = 1, \cdots, n,$$

the law of total probability is proved.

Example 7.1 **A fair coin is flipped. If a head turns up, a fair die is tossed; if a tail turns up, two fair dice are tossed. What is the probability of the event** B **that the sum of the appearing number(s) is equal to 6?**

Solution: **Since** $\mathbb{P}(H) = \frac{1}{2}$ **and** $\mathbb{P}(T) = \frac{1}{2}$, **and**

$$\mathbb{P}(B|H) = \frac{1}{6}$$

$$\mathbb{P}(B|T) = \frac{5}{36}$$

$\{(1,5),(2,4),(3,3),(4,2),(5,1)\}$. **Therefore,**

$$\begin{aligned}\mathbb{P}(B) = \; & \mathbb{P}(H)\mathbb{P}(B|H) + \mathbb{P}(T)\mathbb{P}(B|T) \\ & = \frac{1}{2} \times \frac{1}{6} + \frac{1}{2} \times \frac{5}{36} = 0.15.\end{aligned}$$

Bayes' Rule Given a sequence of mutually exclusive events $S_1, S_2, \cdots, S_n$ **and** $\mathbb{P}(S_i) > 0$. **If event** $A \subset \cup_{i=1}^n S_i$, **then**

$$\mathbb{P}(S_i|A) = \frac{\mathbb{P}(S_i)\mathbb{P}(A|S_i)}{\sum_{k=1}^n \mathbb{P}(S_k)\mathbb{P}(A|S_k)}$$

Proof: According to the definition of conditional probability, we have

$$\mathbb{P}(S_i|A) = \frac{\mathbb{P}(S_i \cap A)}{\mathbb{P}(A)}.$$

Since

$$\mathbb{P}(S_i \cap A) = \mathbb{P}(S_i)\mathbb{P}(A|S_i),$$

and according to the law of total probability,

$$\mathbb{P}(A) = \mathbb{P}(S_1)\mathbb{P}(A|S_1) + \cdots + \mathbb{P}(S_n)\mathbb{P}(A|S_n),$$

this proves the Bayes' rule.

Example 7.2 A man takes either a bus or the subway to work with probabilities 0.3 and 0.7 , respectively. When he takes the bus, he is late 30% of the days. When he takes the subway, he is late 20% of the days.

If the man is late for work on a particular day, what is the probability that he took the bus?

Solution: **Define** $B = \{$ **The man takes a bus** $\}$, $S = \{$ **The man takes the subway** $\}$, **and** $L = \{$ **The man is late on the day** $\}$. **Then,**

$$\mathbb{P}(B) = 0.3, \mathbb{P}(S) = 0.7, \mathbb{P}(L|B) = 0.3, \mathbb{P}(L|S) = 0.2.$$

By Bayes' rule, we have

$$\begin{aligned}\mathbb{P}(B|L) = &\ \frac{\mathbb{P}(B)\mathbb{P}(L|B)}{\mathbb{P}(S)\mathbb{P}(L|S) + \mathbb{P}(B)\mathbb{P}(L|B)} \\ &= \frac{0.3 \times 0.3}{0.7 \times 0.2 + 0.3 \times 0.3} = \frac{0.09}{0.23} = 0.3913.\end{aligned}$$

Example 7.3 **To evaluate the effectiveness of a screening procedure, we will evaluate the probability of a false negative or a false positive using the following notation:**
T^+ : **The test is positive and indicate that the person has the disease.**
T^- : **The test is negative and indicate that the person does not have the disease.**
D^c : **The person really does not have the disease.**
D : **The person really has the disease.**
According to the test results, we found that the sensitivity of the test has following conditional probabilities:

$$\mathbb{P}(T^+|D) = 0.98,$$

and

$$\mathbb{P}(T^-|D^c) = 0.99.$$

If the proportion of the general population infected with this disease is 2 per million, what is
(a) the probability of a false positive,

$$\mathbb{P}(D^c|T^+) \quad ?$$

(b) the probability of a false negative,

$$\mathbb{P}(D|T^-) \quad ?$$

<u>Solution:</u> From the given information, we know the following:

$$\mathbb{P}(D) = 0.000002, \qquad \mathbb{P}(D^c) = 0.999998$$

$$\mathbb{P}(T^+|D) = 0.98 \qquad \mathbb{P}(T^-|D) = 0.02$$

$$\mathbb{P}(T^+|D^c) = 0.01 \qquad \mathbb{P}(T^-|D^c) = 0.99$$

(a) From Bayes' Rule,

$$\begin{aligned}\mathbb{P}(D^c|T^+) &= \frac{\mathbb{P}(D^c \cap T^+)}{\mathbb{P}(T^+)} \\ &= \frac{\mathbb{P}(D^c)\mathbb{P}(T^+|D^c)}{\mathbb{P}(D^c)\mathbb{P}(T^+|D^c) + \mathbb{P}(D)\mathbb{P}(T^+|D)}\end{aligned}$$

Therefore,

$$\begin{aligned}\mathbb{P}(D^c|T^+) &= \frac{0.00999998}{0.01000194} \\ &= 0.999804038\end{aligned}$$

(b) Using a similar calculation,

$$\begin{aligned}\mathbb{P}(D|T^-) &= \frac{\mathbb{P}(D\cap T^-)}{\mathbb{P}(T^-)} \\ &= \frac{\mathbb{P}(D)\mathbb{P}(T^-|D)}{\mathbb{P}(D^c)\mathbb{P}(T^-|D^c)+\mathbb{P}(D)\mathbb{P}(T^-|D)}\end{aligned}$$

Therefore,

$$\begin{aligned}\mathbb{P}(D|T^-) &= \frac{0.00000004}{0.98999806} \\ &= 0.00000004\end{aligned}$$

Hence, the probability of a false positive is near 1 and very likely, while the probability of a false negative is quite small and very unlikely.
<u>Example 7.4</u> If men constitute 47% of the population and tell the truth 78% of the time, while women tell the truth 63% of the time, what is the probability that a person selected at random will answer a question truthfully?

<u>Solution:</u> Define

$$B = \{\textbf{The person interviewd answers truthfully }\}$$

$$A = \{\textbf{ The person interviewed is a man }\}$$

According to the law of total probability, we have

$$\begin{aligned}\mathbb{P}(B) &= \mathbb{P}(A)\mathbb{P}(B|A)+\mathbb{P}(A^c)\mathbb{P}(B|A^c) \\ &= (0.47)(0.78)+(0.53)(0.63)=0.70\end{aligned}$$

Example 7.4 **A worker-operated machine produces a defective item with probability 0.01 if the worker follows the machine's operating instructions exactly, and with probability 0.03 if he does not . If the worker follows the instructions 90% of time, what proportion of all items produced by the machine will be defective? Given that a defective item is produced, what is the conditional probability of the event that the worker exactly follows the machine operating instructions?**

Solution: **Define**
D : **Machine produces a defective item.**
F : **Worker follows instructions.**
Then, we have following information:

$$\mathbb{P}(D|F) = 0.01 \qquad \mathbb{P}(F) = 0.9$$
$$\mathbb{P}(D|F^c) = 0.03 \qquad \mathbb{P}(F^c) = 0.1$$

According to the law of total probability, we have

$$\begin{aligned}\mathbb{P}(D) &= \mathbb{P}(D|F)\mathbb{P}(F) + \mathbb{P}(D|F^c)\mathbb{P}(F^c)\\ &= 0.01(0.9) + 0.03(0.1) = 0.012.\end{aligned}$$

According to the Bayes' Rule, we have

$$\mathbb{P}(F|D) = \frac{\mathbb{P}(F)\mathbb{P}(D|F)}{\mathbb{P}(D)} = \frac{0.9(0.01)}{0.012} = 0.75$$
$$\mathbb{P}(F^c|D) = 0.25$$

Example8.3 Many companies are testing prospective employees for drug use with the intent of improving efficiency and reducing accidents. Suppose a company uses a test that is 98% accuracy to identify a user or a nonuser. To reduce the chance of error, two independent tests are required for each applicant. What are the probabilities of following events?
(a) A nonuser fails both tests.
(b) A drug user is detected.
(c) A drug user is not detected.

Solution: Define
$P_1 := \{$**The first test is positive**$\}$;
$P_2 := \{$**The second test is positive**$\}$;
$N_1 := \{$**The first test is negative**$\}$;
$N_2 := \{$**The second test is negative**$\}$.
$D := \{$ **An applicant is a drug user**$\}$.
$D^c := \{$ **An applicant is not a drug user**$\}$.
We know that $\mathbb{P}(P_1|D) = 0.98$, $\mathbb{P}(P_2|D) = 0.98$, $\mathbb{P}(N_1|D^c) = 0.98$ **and** $\mathbb{P}(N_2|D^c) = 0.98$. **Then,** $\mathbb{P}(N_1|D) = 1 - 0.98 = 0.02$, $\mathbb{P}(N_2|D) = 0.02$, $\mathbb{P}(P_1|D^c) = 0.02$ **and** $\mathbb{P}(P_2|D^c) = 0.02$.
(a) $\mathbb{P}(P_1 \cap P_2|D^c) = \mathbb{P}(P_1|D^c)\mathbb{P}(P_2|D^c) = (0.02)(0.02) = 0.0004$

(b)

$$
\begin{aligned}
&\mathbb{P}((P_1 \cap N_2) \cup (N_1 \cap P_2) \cup (P_1 \cap P_2)|D) \\
&= \mathbb{P}((P_1 \cap N_2)|D) + \mathbb{P}((N_1 \cap P_2)|D) + \mathbb{P}((P_1 \cap P_2)|D) \\
&= \mathbb{P}(P_1|D)\mathbb{P}(N_2|D) + \mathbb{P}(N_1|D)\mathbb{P}(P_2|D) + \mathbb{P}(P_1|D)\mathbb{P}(P_2|D) \\
&= (0.98)(0.02) + (0.02)(0.98) + (0.98)(0.98) = 0.9996
\end{aligned}
$$

(c)

$$\mathbb{P}(N_1 \cap N_2|D) = \mathbb{P}(N_1|D)\mathbb{P}(N_2|D) = (0.02)(0.02) = 0.0004$$

Example As items come to the end of a production line, an inspector chooses items to undergo a complete inspection. Of all items produced 20% are defective. 50% of all defective items go through a complete inspection, and 30% of all good items go through a complete inspection. Given that an item is completely inspected, what is the probability that it is defective?

Solution: Define
D : An item defective.
C : An item is completely inspected.
Then, we have following information:

$$\mathbb{P}(C|D) = 50\% \qquad \mathbb{P}(D) = 20\%$$

$$\mathbb{P}(C|D^c) = 30\% \qquad \mathbb{P}(D^c) = 80\%$$

According to the Bayes' Rule, we have

$$\mathbb{P}(D|C) = \frac{\mathbb{P}(D)\mathbb{P}(C|D)}{\mathbb{P}(D)\mathbb{P}(C|D) + \mathbb{P}(D^c)\mathbb{P}(C|D^c)} = \frac{0.2(0.5)}{0.2(0.5) + 0.8(0.3)} = \frac{10}{34}$$

Discrete Random Variables and Their Distribution

In order to use the existing mathematical tools to find out the probabilities of events, mean, variance, and so on, statisticians figured out a nice way as follows:
We associate each simple event with a real number. For

example, in the coin tossing experiment, we define:

$$\begin{array}{ll} \Lambda = \{ & 1, \quad\quad 0 \ \} \subset \mathbb{R} = \{\textbf{all the real numbers}\} \\ & \uparrow \quad\quad \uparrow \\ \Omega = \{ & \textbf{head, tail} \ \} \end{array}$$

This is called a transformation. Once transformed, we can use the existing mathematical tools. For our convenience, in this section we denote the sample space by $\Omega = \{\omega_1, \omega_2, \cdots, \omega_n\}$**, in which each** ω_i **is a simple event.**

<u>Definition 9.1</u>
Consider a random experiment with sample space $\Omega = \{\omega_1, \omega_2, \cdots, \omega_n\}$**. A function** $X : \ \Omega \to \mathbb{R}$ **which assigns to each simple event** $\omega \in \Omega$ **<u>one and only one</u> real number** $X(\omega) = x \in \mathbb{R}$**, is called a <u>random variable</u>.**
$\Lambda := \{x_i : x_i = X(\omega_i), \omega_i \in \Omega\}$ **is called the <u>range space</u> of the random variable (r.v.)** X**.**

<u>Definition 9.2</u>
If Λ **is a finite set or a countable set, then** X **is called a <u>discrete</u> <u>random variable</u>. Given a discrete random variable** X **with range space** $\Lambda = \{x_1, \cdots, x_n\}$**. For each** $x_i \in \Lambda$**, we define**

$$p(x_i) = \mathbb{P}(\{\omega : X(\omega) = x_i\}).$$

Then, $p(x)$ **is called the <u>discrete density function</u> of r.v.** X **and the following table is called a <u>discrete probability distribution</u> of r.v.** X**.**

The distribution table of r.v X

x	x_1	x_2	$\cdots$	x_n
$p(x)$	$p(x_1)$	$p(x_2)$	$\cdots$	$p(x_n)$

Example 9.1 Consider a random experiment that consists of tossing two fair coins successively and let X be equal to the number of heads observed. Find its sample space, range space, and distribution.

Solution: The sample space $\Omega = \{\omega_1 = \{TT\}, \omega_2 = \{HT\}, \omega_3 = \{TH\}, \omega_4 = \{HH\}\}$ **and** $\Lambda := \{x_1 = 0, x_2 = 1, x_3 = 2\}$**.**

$$p(x_1) = p(0) = \mathbb{P}(X = 0) = \mathbb{P}(\{TT\}) = \frac{1}{4},$$

$$p(x_2) = p(1) = \mathbb{P}(X = 1) = \mathbb{P}(\{TH, HT\}) = \mathbb{P}(\{TH\}) + \mathbb{P}(\{HT\}) = \frac{2}{4},$$

$$p(x_3) = p(2) = \mathbb{P}(X = 2) = \mathbb{P}(\{HH\}) = \frac{1}{4}.$$

Therefore,

The distribution table of r.v X

x	**0**	**1**	**2**
$p(x)$	$1/4$	$1/2$	$1/4$

Remark:

(a) $0 \leq p(x_i) \leq 1$**,**

(b) $\sum_{x_i \in \Lambda} p(x_i) = \mathbb{P}(\Omega) = 1$

<u>**Definition 9.3**</u> **Let** X **be a discrete r.v. with range space** $\Lambda = \{x_1, \cdots, x_n\}$ **and distribution** $p(x_i), i = 1, \cdots, n$**. The** <u>**expected value**</u> **or** <u>**mean**</u> **of** X **is defined by**

$$\mu = \mathbb{E}(X) = \sum_{x_i \in \Lambda} x_i p(x_i).$$

The <u>**variance**</u> **of** X **is defined by**

$$\sigma^2 = \mathbb{E}[(X-\mu)^2] = \sum_{x_i \in \Lambda} (x_i - \mu)^2 p(x_i).$$

The <u>**standard deviation**</u> **of** X **is defined as** $\sigma = \sqrt{\sigma^2}$**.**

<u>**Example 9.2**</u> **A company has five applicants for two positions: two women and three men. Suppose that the five applicants are equally qualified and no preference is given for choosing either gender. Let** X **equal the number of women chosen to fill the two positions.**
(a) Find the probability distribution of X**;**
(b) Find the mean and standard deviation of X**.**

<u>**Example 9.2**</u> **A company has five applicants for two positions: two women and three men. Suppose that the five applicants are equally qualified and no preference is given for choosing either gender. Let** X **equal the number of women chosen to fill the two positions.**
(a) Find the probability distribution of X**;**
(b) Find the mean and standard deviation of X**.**

<u>Solution:</u> (a). Define short notations as follows:
M_1**: Candidate man one;**
M_2**: Candidate man two;**
M_3**: Candidate man three;**
W_1**: Candidate woman one;**
W_2**: Candidate woman two.**
The sample space

$$\Omega = \{\omega_1 = \{M_1M_2\}, \omega_2 = \{M_1M_3\}, \omega_3 = \{M_2M_3\},$$

$$\omega_4 = \{W_1M_1\}, \omega_5 = \{W_1M_2\}, \omega_6 = \{W_1M_3\}, \omega_7 = \{W_2M_1\},$$

$$\omega_8 = \{W_2M_2\}, \omega_9 = \{W_2M_3\}, \omega_{10} = \{W_1W_2\}\}$$

and the range space

$$\Lambda := \{x_1 = 0, x_2 = 1, x_3 = 2\}.$$

$$p(x_1) = p(0) = \mathbb{P}(X = 0)$$

$$= \mathbb{P}(\{\omega_1, \omega_2, \omega_3\})$$

$$= \mathbb{P}(\{\{M_1M_2\}, \{M_1M_3\}, \{M_2M_3\}\})$$

$$= \tfrac{3}{10},$$

$$p(x_2) = p(1) = \mathbb{P}(X = 1) = \mathbb{P}(\{\omega_4, \omega_5, \omega_6, \omega_7, \omega_8, \omega_9\})$$

$$= \mathbb{P}(\{\{W_1M_1\}, \{W_1M_2\}, \{W_1M_3\}, \{W_2M_1\}, \{W_2M_2\}, \{W_2M_3\}\})$$

$$= \tfrac{6}{10},$$

$$p(x_3) = p(2) = \mathbb{P}(X = 2) = \mathbb{P}(\{\omega_{10} = \{W_1W_2\}\}) = \tfrac{1}{10}.$$

Therefore,

The distribution table of r.v X

x	0	1	2
$p(x)$	3/10	6/10	1/10

(b)

$$\mu = \mathbb{E}(X) = \textstyle\sum_{x_i\in\Lambda} x_i p(x_i)$$

$$= 0(3/10) + 1(6/10) + 2(1/10) = 8/10.$$

The variance of X is

$$\sigma^2 = \mathbb{E}[(X-\mu)^2] = \textstyle\sum_{x_i\in\Lambda}(x_i-\mu)^2 p(x_i)$$

$$= (0-8/10)^2(3/10) + (1-8/10)^2(6/10)+$$

$$(2-8/10)^2(1/10) = 0.36$$

and

$$\sigma = \sqrt{\sigma^2} = 0.6$$

<u>Example 9.3</u> A jar contains four coins: a nickel, a dime, a quarter, and a half-dollar. Three coins are randomly selected from the jar. Let X be equal to the total amount drawn.

a. List all the simple events in the sample space Ω and find the range space Λ.

b. Find the probability $\mathbb{P}(X \geq 0.5$ dollar$)$.

c. Find the probability distribution of X.

<u>Solution:</u>
a. Denote:
N: nickel;
D: dime;
Q: quarter;
H: half-dollar.
and $E_1 = (NDQ)$,$E_2 = (NDH)$, $E_3 = (NQH)$, $E_4 = (DQH)$.
Then, $\Omega = \{E_1, E_2, E_3, E_4\}$ **and** $\Lambda := \{x_1 = 0.4, x_2 = 0.65, x_3 = 0.80, x_4 = 0.85\}$.
b. The simple event along with their monetary values follow:

$$E_1 = NDQ = \$0.4$$
$$E_2 = NDH = \$0.65$$
$$E_3 = NQH = \$0.80$$
$$E_4 = DQH = \$0.85$$

$$\mathbb{P}(X \geq 0.5 \textbf{ dollar}) = \mathbb{P}(E_2) + \mathbb{P}(E_3) + \mathbb{P}(E_4) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}$$

c. We have

$$p(x_1) = p(0.4) = \mathbb{P}(X = 0.4) = \mathbb{P}(\{NDQ\}) = \frac{1}{4},$$
$$p(x_2) = p(0.65) = \mathbb{P}(X = 0.65) = \mathbb{P}(\{NDH\}) = \frac{1}{4},$$
$$p(x_3) = p(0.80) = \mathbb{P}(X = 0.80) = \mathbb{P}(\{NQH\}) = \frac{1}{4}.$$

$$p(x_4) = p(0.85) = \mathbb{P}(X = 0.85) = \mathbb{P}(\{DQH\}) = \frac{1}{4}.$$

Therefore,

The distribution table of r.v X

x	**0.4**	**0.65**	**0.80**	**0.85**
$p(x)$	$1/4$	$1/4$	$1/4$	$1/4$

Example 9.4 A student prepares for a quiz by studying a list of ten problems. She only can solve six of them. For the quiz, the instructor selects five questions at random from the list of ten. Let X be the number of questions she can solve.
a Find total number of simple events in the sample space and the range space.
b Find the probability distribution of X.

Solution:
a. The total number of simple events is the sample space is equal to $C_5^{10} = 252$. Then, $\Lambda := \{x_1 = 1, x_2 = 2, x_3 = 3, x_4 = 4, x_5 = 5\}$.

b. We have

$$p(x_1) = p(1) = \mathbb{P}(X = 1) == \frac{C_1^6 C_4^4}{252},$$

$$p(x_2) = p(2) = \mathbb{P}(X = 2) == \frac{C_2^6 C_3^4}{252},$$

$$p(x_3) = p(3) = \mathbb{P}(X = 3) == \frac{C_3^6 C_2^4}{252},$$
$$p(x_4) = p(4) = \mathbb{P}(X = 4) == \frac{C_4^6 C_1^4}{252},$$
$$p(x_5) = p(5) = \mathbb{P}(X = 5) == \frac{C_5^6 C_0^4}{252},$$

Therefore,

The distribution table of r.v X

x	**1**	**2**	**3**	**4**	**5**
$p(x)$	$\frac{C_1^6 C_4^4}{252}$	$\frac{C_2^6 C_3^4}{252}$	$\frac{C_3^6 C_2^4}{252}$	$\frac{C_4^6 C_1^4}{252}$	$\frac{C_5^6 C_0^4}{252}$

Example 9.5

In a pocket there are 3 black and 2 white balls. Balls are identical except their colors. We randomly draw a ball and observe its color two times successively, with replacement. We define that X is the number of black balls observed.

a Find the sample space and the range space.

b Find the probability $\mathbb{P}(X \geq 1)$.

c Find the probability distribution of X.

Solution:

a. Denote:

B: observe a black ball;

W: observe a white ball;

and

$$E_1 = (BB), \quad E_2 = (BW), \quad E_3 = (WB), \quad E_4 = (WW),$$

Then, $\Omega = \{E_1, E_2, E_3, E_4\}$ **and** $\Lambda := \{x_1 = 0, x_2 = 1, x_3 = 2\}$**.**
b.

$$\begin{aligned}\mathbb{P}(X \geq 1) &= 1 - \mathbb{P}(X = 0) = 1 - \mathbb{P}(\{WW\}) = 1 - \mathbb{P}(W)\mathbb{P}(W) \\ &= 1 - (\frac{2}{5})(\frac{2}{5}) = 1 - \frac{4}{25} = \frac{21}{25}\end{aligned}$$

c. We have

$$p(x_1) = p(0) = \mathbb{P}(X = 0) = \mathbb{P}(\{WW\}) = \frac{4}{25},$$

$$\begin{aligned}&p(x_2) = p(1) = \mathbb{P}(X = 1) = \mathbb{P}(\{BW\} \cup \{WB\}) \\ &= \mathbb{P}(\{BW\}) + \mathbb{P}(\{WB\}) \\ &= 2\mathbb{P}(\{W\})\mathbb{P}(\{B\}) = 2(\frac{2}{5})(\frac{3}{5}) = \frac{12}{25},\end{aligned}$$

$$\begin{aligned}&p(x_3) = p(2) = \mathbb{P}(X = 2) = \mathbb{P}(\{BB\}) \\ &= \mathbb{P}(\{B\})\mathbb{P}(\{B\}) = (\frac{3}{5})(\frac{3}{5}) = \frac{9}{25},\end{aligned}$$

Therefore,

The distribution table of r.v X

x	**0**	**1**	**2**
$p(x)$	4/25	12/25	9/25

The Binomial Probability Distribution

Definition 10.1 **A binomial experiment is one that has following three characteristics:**

(1) The experiment consists of n **ordered, independent,**

identical trials.
(2) Each trial has two possible outcomes: success A, failure $\bar{A}$.
(3) $0 < \mathbb{P}(A) = p < 1$ **and** $\mathbb{P}(\bar{A}) = 1 - p = q$**.**
Remark: In binomial experiment, if $p \neq q$, then simple events are not equally likely. Therefore, it is not a classical probability model.
Example10.1 Consider a binomial experiment of flipping a biased coin three times successively. Let A be the event of observing a head and $\bar{A}$ be the event of observing a tail.
Suppose that

$$\mathbb{P}(A) = \frac{1}{3} \qquad \mathbb{P}(\bar{A}) = \frac{2}{3}$$

Then

$$\Omega = \{AAA, \bar{A}AA, A\bar{A}A, AA\bar{A}, \bar{A}\bar{A}A, \bar{A}A\bar{A}, A\bar{A}\bar{A}, \bar{A}\bar{A}\bar{A}\}$$

$$\mathbb{P}(AAA) = \mathbb{P}(A)\mathbb{P}(A)\mathbb{P}(A) = \frac{1}{27}$$

$$\mathbb{P}(A\bar{A}A) = \mathbb{P}(A)\mathbb{P}(\bar{A})\mathbb{P}(A) = (\frac{1}{3})(\frac{2}{3})(\frac{1}{3}) = \frac{2}{27}$$

Therefore, it is not a classical probability model.
Let X be number of heads observed. Find the probability $\mathbb{P}(X = 2)$.

$$\mathbb{P}(X = 2) = \mathbb{P}(\{AA\bar{A}\}, \{A\bar{A}A\}, \{\bar{A}AA\})$$

$$\mathbb{P}(\{AA\bar{A}\}) = \mathbb{P}(A)\mathbb{P}(A)\mathbb{P}(\bar{A}) = ppq$$

$$\mathbb{P}(\{\bar{A}AA\}) = \mathbb{P}(A)\mathbb{P}(A)\mathbb{P}(\bar{A}) = ppq$$
$$\mathbb{P}(\{A\bar{A}A\}) = \mathbb{P}(A)\mathbb{P}(A)\mathbb{P}(\bar{A}) = ppq$$
$$\mathbb{P}(X = 2) = C_2^3 ppq = C_2^3 p^2 q^{3-2}$$

Generally, consider a binomial experiment with n trials. let X be the number of successes in the n trials. What is the probability of event $\{X = k\}$. Then,

$$\mathbb{P}(X = k) = C_k^n p^k q^{n-k}$$

(For C_k^n, Consider there are n positions and how many ways to choose $n - k$ positions to put bars on tops of these positions.)

Definition 10.2 **A binomial experiment consists of n ordered, independent, identical trials with probability of success p and probability of failure $q = 1 - p$ on each trial. Let X be the number of successes in the n trials. Then,**

The distribution table of r.v X

x	0	1	2	$\cdots$	n
$p(x)$	$C_0^n p^0 q^n$	$C_1^n p^1 q^{n-1}$	$C_2^n p^2 q^{n-2}$	$\cdots$	$C_n^n p^n q^0$

is called the binomial distribution. Its mean $\mu = np$, its variance $\sigma^2 = npq$, its standard deviation $\sigma = \sqrt{npq}$

Example 9.6

In a pocket there are 3 black and 2 white balls. Balls

are identical except their colors. We randomly draw a ball and observe its color three times successively, with replacement. We define that X is the number of black balls observed.
a Find the sample space and the range space.
b Find the probability $\mathbb{P}(X \geq 1)$.
c Find the probability distribution of X.

Solution:
a. Denote:
B: observe a black ball;
W: observe a white ball;
and

$$E_1 = (BBB), \quad E_2 = (BBW), \quad E_3 = (BWB),$$
$$E_4 = (BWW), E_5 = (WBB), \quad E_6 = (WBW),$$
$$E_7 = (WWB), \quad E_8 = (WWW).$$

Then, $\Omega = \{E_1, E_2, E_3, E_4, E_5, E_6, E_7, E_8\}$ **and** $\Lambda := \{x_1 = 0, x_2 = 1, x_3 = 2, x_4 = 3\}$**.**
b.

$$\begin{aligned} \mathbb{P}(X \geq 1) &= 1 - \mathbb{P}(X = 0) = 1 - \mathbb{P}(\{WWW\}) \qquad (0.1) \\ &= 1 - \mathbb{P}(W)\mathbb{P}(W)\mathbb{P}(W) \\ &= 1 - (\frac{2}{5})(\frac{2}{5})(\frac{2}{5}) = 1 - \frac{8}{125} = \frac{117}{125} \end{aligned}$$

c. We have

$$p(x_1) = p(0) = \mathbb{P}(X = 0) = \mathbb{P}(\{WWW\}) = \frac{8}{125},$$

$$p(x_2) = p(1) = \mathbb{P}(X = 1)$$
$$= \mathbb{P}(\{BWW\} \cup \{WBW\} \cup \{WWB\})$$
$$= \mathbb{P}(\{BWW\}) + \mathbb{P}(\{WBW\}) + \mathbb{P}(\{WWB\})$$
$$= 3\mathbb{P}(\{W\})\mathbb{P}(\{W\})\mathbb{P}(\{B\}) = 3(\frac{2}{5})(\frac{2}{5})(\frac{3}{5}) = \frac{36}{125},$$

$$p(x_3) = p(2) = \mathbb{P}(X = 2) = \mathbb{P}(\{WBB\} \cup \{BWB\} \cup \{BBW\})$$
$$= \mathbb{P}(\{WBB\}) + \mathbb{P}(\{BWB\}) + \mathbb{P}(\{BBW\})$$
$$= 3\mathbb{P}(\{B\})\mathbb{P}(\{B\})\mathbb{P}(\{W\}) = 3(\frac{3}{5})(\frac{3}{5})(\frac{2}{5}) = \frac{54}{125},$$

$$p(x_4) = p(3) = \mathbb{P}(X = 3) = \mathbb{P}(\{BBB\}) = \mathbb{P}(\{B\})\mathbb{P}(\{B\})\mathbb{P}(\{B\})$$
$$= (\frac{3}{5})(\frac{3}{5})(\frac{3}{5}) = \frac{27}{125}.$$

Therefore,

The distribution table of r.v X

x	**0**	**1**	**2**	**3**
$p(x)$	8/125	36/125	54/125	27/125

Example 10.1
Suppose that a family will certainly have 5 children, but each child being a boy or girl is totally uncertain with equal probability. Let X be the number of boys this family will have. Find the distribution of X and the probability $\mathbb{P}(X \geq 1)$.

Example10.2 Consider a binomial experiment of flipping a biased coin twenty times successively. Let A be

the event of observing a head and $\bar{A}$ be the event of observing a tail.
Suppose that

$$\mathbb{P}(A) = \frac{1}{5} \qquad \mathbb{P}(\bar{A}) = \frac{4}{5}$$

Let X be number of heads observed in the twenty flippings. (a) Find the probability $\mathbb{P}(X = 8)$. (b) Find its mean and variance. (c) Find the probability of event $(X > 2)$.

<u>solution:</u> **(a) According to the binomial distribution formula, we have**

$$\mathbb{P}(X = 8) = C_8^{20}(\frac{1}{5})^8(\frac{4}{5})^{20-8}.$$

We have $\mathbb{P}(X \le 8) = 0.990$ and $\mathbb{P}(X \le 7) = 0.968$ according to the table 1 ($n = 20$, $k = 8, k = 7$, $p = 0.2$) on page 684.
$\mathbb{P}(X = 8) = \mathbb{P}(X \le 8) - \mathbb{P}(X \le 7) = 0.022$ **.**
(b) Its mean is equal to

$$\mu = 20(1/5) = 4$$

and its variance is equal to

$$\sigma^2 = 20(1/5)(4/5) = \frac{16}{5}$$

(c) $\mathbb{P}(X > 2) = 1 - \mathbb{P}(X \le 2) = 1 - 0.206 = 0.794$
<u>**Example10.3**</u> **Consider an experiment of randomly drawing a chip 8 times successively with replacement from a box containing two black chips and three white chips.**

The chips are identical except their colors. Let X be the number of black chips observed in the 8 drawings. (a) Find the probability $\mathbb{P}(X = 5)$. (b) Find its mean and variance. (c) Find the probability of event $(X > 2)$.

<u>solution:</u> **(a) According to the binomial distribution formula, we have**

$$\mathbb{P}(X = 5) = C_5^8 (\frac{2}{5})^5 (\frac{3}{5})^{8-5}.$$

We have $\mathbb{P}(X \leq 5) = 0.950$ and $\mathbb{P}(X \leq 4) = 0.826$ according to the table 1 ($n = 8$, $k = 5, k = 4$, $p = 0.4$) on page 681. $\mathbb{P}(X = 5) = \mathbb{P}(X \leq 5) - \mathbb{P}(X \leq 4) = 0.124$ **.**
(b) Its mean is equal to

$$\mu = 8(2/5) = 16/5$$

and its variance is equal to

$$\sigma^2 = 8(2/5)(3/5) = \frac{48}{25}$$

(c) $\mathbb{P}(X > 2) = 1 - \mathbb{P}(X \leq 2) = 1 - 0.315 = 0.685$
<u>**Example**</u>**: A factory employs several thousand workers, of whom 30% are Hispanic. If the 15 members of the union executive committee were chosen from the workers at random,let X be the number of Hispanics on the committee.**
(a) Find $\mathbb{P}(X = 3)$.
(b) Find $\mathbb{P}(X \leq 3)$.

<u>Example</u>: Suppose that early statewide election returns indicate totals of $33,000$ votes for candidate A versus $27,000$ for candidate B, and that these early returns can be regarded as a random sample selected from the population of all $10,000,000$ eligible voters in the state. Let X be the number of votes for A.
(a) If the statewide vote will be split $50-50$, find the expected number of votes for A in the sample of $60,000$ early returns.
(b) Find the standard deviation of X.
(c) Find the probability that $(X > 40,000)$.
<u>Solution</u>:
(a) $\mu = np = 60,000 \times 0.5 = 30,000.$
(b) $\sigma = \sqrt{npq} = \sqrt{15,000} = 122.47449$
(c) $\mathbb{P}(X > 40,000) = 1 - \sum_{i=0}^{40,000} C_i^{60,000} (0.5)^i (0.5)^{60,000-i}$