

## 進佳數學團隊 – Dr. Herbert Lam 林康榮博士

### HKDSE Mathematics M1

## II. THE BINOMIAL DISTRIBUTION

---

### 1. Bernoulli distribution

A Bernoulli experiment results in any one of two possible outcomes, which are often classified as “success” and “failure”. For example, the inspection of a lot of items is associated with the sample space

$$S = \{D, ND\}$$

in which  $D$  denotes the event of getting a defective and

$ND$  denotes the event of getting a nondefective.

If one of them is termed “success”, then the other is termed “failure”.

Some experiments look like Bernoulli, but not really. For instance, John is playing a game of chess with his friend. One might associate “success” with “winning the game” and “failure” with “losing the game”. However, there is also another possible outcome, a “tie”!

Let us quantify the event “success” by 1 and “failure” by 0.

Then  $Y = \{0, 1\}$  is called a Bernoulli random variable. Suppose the probability of a “success” is  $p$ . Then the probability function of  $Y$  is given by

$$P(Y = 1) = p, \quad P(Y = 0) = 1 - p.$$

### 2. Basic features of a binomial experiment

Let  $Y_1, Y_2, \dots, Y_n$  be  $n$  independent Bernoulli random variables, and

let  $X = Y_1 + Y_2 + \dots + Y_n$ .

Then  $X$  is called a binomial random variable, which takes on the values  $0, 1, 2, \dots, n$ .

This approach is rather mathematical. In fact, we can commence our study of the binomial distribution intuitively as follows.

A binomial experiment possesses the following properties:

1. There are  $n$  identical observations or trials.
2. Each trial has two possible outcomes, one called success and the other failure. The outcomes are mutually exclusive and collectively exhaustive.
3. The probabilities of success  $p$  and of failure  $1-p$  remain the same for all trials
4. The outcomes of trials are independent of each other.

## 2. The probability function

In a binomial experiment with a constant probability  $p$  of success in each trial, the probability distribution of the binomial random variable  $X$ , the number of successes in  $n$  independent trials, is called the **binomial distribution**.

Since the  $n$  trials result in  $x$  successes and  $n-x$  failures, the probability is  $p^x(1-p)^{n-x}$ . Moreover, there are  $C_x^n$  combinations of successes and failures. Therefore, the probability function is given by

$$P(X = x) = C_x^n p^x (1-p)^{n-x}, \quad x = 0, 1, 2, \dots, n.$$

The notation  $X \sim b(n, p)$  is used to describe a binomial random variable  $X$  with parameters  $n$  and  $p$ . Furthermore, the mean and variance of the above binomial distribution are  $np$  and  $np(1-p)$  respectively.

### Example 1

A trading company has four telephone lines. Suppose the probability that any one of the lines is busy at an instant is  $\frac{1}{3}$ .

- (a) Calculate the probability that
  - (i) two of the four lines are busy.
  - (ii) at least one of the four lines is busy.
  - (iii) an incoming phone call cannot be answered immediately.
- (b) Suppose at an instant one staff member is on the telephone. What is the probability that another two lines are busy?

 **[Solution]**

(a) Let  $X$  denote the number of telephone lines that are busy.

$$\begin{aligned} \text{(i)} \quad P(X = 2) &= C_2^4 \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^2 \\ &= \frac{4 \cdot 3}{2} \cdot \frac{1}{9} \cdot \frac{4}{9} = \frac{8}{27}. \end{aligned}$$

$$\begin{aligned} \text{(ii)} \quad P(X \geq 1) &= 1 - P(X = 0) = 1 - C_0^4 \left(\frac{1}{3}\right)^0 \left(\frac{2}{3}\right)^4 \\ &= 1 - \frac{16}{81} = \frac{65}{81}. \end{aligned}$$

(iii) An incoming call not answered immediately implies all the four lines are busy.

$$P(X = 4) = \left(\frac{1}{3}\right)^4 = \frac{1}{81}.$$

(b) Since it is given that one telephone line is engaged, the outcome of  $X = 0$  is removed. We have the conditional probability

$$\begin{aligned} &P(3 \text{ lines are busy/at least 1 line is busy}) \\ &= \frac{P(X = 3 \ \& \ X \geq 1)}{P(X \geq 1)} = \frac{P(X = 3)}{P(X \geq 1)} = \frac{C_3^4 \left(\frac{1}{3}\right)^3 \left(\frac{2}{3}\right)}{\frac{65}{81}} = \frac{4 \times 2}{65} = \frac{8}{65}. \end{aligned}$$

### Example 2

Machine A and B turn out, respectively, 10% and 90% of the total production of a certain article. Suppose the probability that machine A turns out a defective article is 0.01 and that machine B turns out a defective one is 0.05.

- What is the probability that an article taken at random from a day's production was nondefective?
- In a quality control process 10 articles from a batch are randomly selected for inspection. Acceptance of the batch allows no more than one defective. Find the probability that the batch is rejected.
- A batch was rejected yesterday because two defective articles are found. What is the probability that both items were manufactured by
  - the same machine?
  - different machines?
- State the relevant assumptions required for (b) and (c).

 **[Solution]**

- (a) For a randomly selected article, define

 $A$ : the article is manufactured by machine A $B$ : the article is manufactured by machine B $ND$ : the article is nondefective $D$ : the article is defective.

Then we have

$$P(A) = 0.1, \quad P(B) = 0.9$$

$$P(ND/A) = 0.99, \quad P(ND/B) = 0.95$$

$$P(D/A) = 0.01, \quad P(D/B) = 0.05.$$

By the theorem of total probability on P. 68,

$$P(ND) = P(A)P(ND/A) + P(B)P(ND/B)$$

$$= (.1)(.99) + (.9)(.95) = .954.$$

$$P(D) = 1 - P(ND) = 1 - .954 = .046.$$

- (b) Let
- $X$
- be the number of defective articles found in the inspection process. Then
- $X \sim b(10, 0.046)$
- .

P(the batch is rejected)

$$= 1 - P(X \leq 1) = 1 - P(X = 0) - P(X = 1)$$

$$= 1 - C_0^{10} (.046)^0 (.954)^{10} - C_1^{10} (.046)(.954)^9 = .0745.$$

- (c) By Bayes' theorem,

$$\begin{aligned} P(A/D) &= \frac{P(A)P(D/A)}{P(A)P(D/A) + P(B)P(D/B)} \\ &= \frac{(.1)(.01)}{(.1)(.01) + (.9)(.05)} = \frac{1}{1 + 45} = \frac{1}{46}. \end{aligned}$$

As an article is either manufactured by machine A or B, it follows that,

$$P(B/D) = 1 - P(A/D) = \frac{45}{46}.$$

- (i) P(both articles manufactured by the same machine)  
 $= P(\text{both manufactured by A}) + P(\text{both manufactured by B})$   
 $= P(A/D)^2 + P(B/D)^2$   
 $= \left(\frac{1}{46}\right)^2 + \left(\frac{45}{46}\right)^2 = \frac{2026}{2116} = .9575.$
- (ii) P(both articles manufactured by different machines)  
 $= 1 - P(\text{both articles manufactured by the same machine})$   
 $= 1 - .9575 = .0425.$

Alternatively, this probability can also be calculated by

$$P(A/D) \cdot P(B/D) + P(B/D) \cdot P(A/D) = 2 \cdot \frac{1}{46} \cdot \frac{45}{46}, \text{ or}$$

P(1 defective article manufactured by machine A)

$$= C_1^2 \left( \frac{1}{46} \right) \left( 1 - \frac{1}{46} \right).$$

- (d) The binomial assumptions are required, namely,
- (i) independent trials and
  - (i) the probability of drawing an defective article is constant.
- In fact, sampling without replacement does not meet the two requirements. However, the batch size is very large and therefore (i) & (ii) are considered satisfied.  
See Example 3 for detail.

### Example 3

A machine produces, on the average, 5% of defective parts. If 10 parts are selected at random from a lot of size 1000 for inspection, what is the probability that exactly three will be defective?

#### [Solution]

Of the 1000 parts in the lot, there are 50 defective ones.

P(exactly 3 will be defective)

$$\begin{aligned} &= \frac{C_3^{50} \cdot C_7^{950}}{C_{10}^{1000}} = \frac{50! \cdot 950!}{3!47! \cdot 7!943!} \\ &= \frac{50 \times 49 \times 48 \times 47! \cdot 950 \times \dots \times 944 \times 943!}{1000 \times 999 \times \dots \times 991 \times 990!} \\ &= \frac{10!}{3!7!} \underbrace{\left( \frac{950}{1000} \times \frac{949}{999} \times \dots \times \frac{944}{994} \right)}_{\approx (.97)^7} \underbrace{\left( \frac{50}{993} \times \frac{49}{992} \times \frac{48}{991} \right)}_{\approx (.05)^3} \\ &\approx C_3^{10} (.05)^3 (.95)^7. \end{aligned}$$

**Example 4**

A machine which is powered by three similar electrical devices will function properly if at least two of these devices are serviceable. Experience indicates that the probability of any device failing in less than 50 hours is 0.2, while failing in less than 100 hours is 0.6.

Find the probability that the machine will function properly

- (a) for more than 50 hours,  
 (b) between 50 and 100 hours.

 **[Solution]**

Let  $X$  denote the number of devices failing in less than 50 hours,

$Y$  denote the number of devices failing in less than 100 hours.

We are given that

$$X \sim b(3, 0.2), \text{ i.e. } P(X = x) = C_x^3 (.2)^x (.8)^{3-x}$$

$$Y \sim b(3, 0.6), \text{ i.e. } P(Y = y) = C_y^3 (.6)^y (.4)^{3-y}.$$

Define the events

$A$ : the machine functions for more than 50 hours

$B$ : the machine functions for more than 100 hours.

- (a)  $P(A) = P(\text{at most one device fails in 50 hours})$

$$= P(X \leq 1) = P(X = 0) + P(X = 1)$$

$$= C_0^3 (.2)^0 (.8)^3 + C_1^3 (.2)(.8)^2$$

$$= .512 + .384 = .896.$$

- (b)  $P(B) = P(\text{at most one device fails in 100 hours})$

$$= P(Y \leq 1) = P(Y = 0) + P(Y = 1)$$

$$= C_0^3 (.6)^0 (.4)^3 + C_1^3 (.6)(.4)^2$$

$$= .064 + .288 = .352.$$

Note that  $B \subset A$ . Then  $A - B$  or  $A \cap \bar{B}$  is the event  
 “the machine functions between 50 and 100 hours”

and we have

$$P(A - B) = P(A) - P(B) = .896 - .352 = .544.$$

**Note:**

As time is a continuous variable, the terms “more than 50 hours” and “at least 50 hours” have the same meaning. Also, “between 50 and 100 hours is well defined. It doesn’t matter if the end points 50 or 100 are inclusive.

### III. THE POISSON DISTRIBUTION

---

#### 1. Basic features

Experiments yielding numerical values of a random variable  $X$ , the number of successes (observations) occurring during a given time interval (or in a specified region) are often called Poisson experiments.

A **Poisson experiment** has the following properties:

1. The number of successes in any interval is independent of the number of successes in other interval.
2. The probability of a single success occurring during a short interval is proportional to the length of the time interval and does not depend on the number of successes occurring outside this time interval.
3. The probability of more than one success in a very small interval is negligible.

Some well known examples of Poisson experiment are

1. The number of customers who arrive during a time period of length  $t$ ,
2. The number of telephone calls per hour received by an office,
3. The number of typing errors per page,
4. The number of accidents per day at a junction.

#### 2. The probability function

The probability distribution of the Poisson random variable  $X$  is called the **Poisson distribution**. The probability function is

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, 2, \dots$$

The mean and variance of the above Poisson distribution are  $\lambda$ . This is also known as the average number of “successes” occurring in the given time interval. Note that  $\lambda$  is the only parameter that appears in the probability function. In other words, a Poisson distribution is completely determined if  $\lambda$  is known.

### Example 1

Albert is a life insurance agent. Assume he makes, on the average, one sale per week, and the number of sales behaves close to a Poisson distribution.

- (a) What is the probability that Albert makes
- exactly 3 sales in a two-week period?
  - at least 3 sales in a three-week period?
- (b) What is the probability that he will make only one sale in the coming December?
- (c) Albert has received e-mail confirmation from two of his friends that each of them will purchase a life insurance policy from him next month.

What is the probability that he will make at least 4 sales next month?

(Take 1 week = 7 days and 1 month = 30 days)

### [Solution]

- (a) (i)  $\lambda = 2/2$ -week.

$$P(X \geq 3) = \frac{e^{-2} 2^3}{3!} = .1804.$$

- (ii)  $\lambda = 3/3$ -week.

$$\begin{aligned} P(X \geq 3) &= 1 - P(X < 3) \\ &= 1 - \sum_{x=0}^2 \frac{e^{-3} 3^x}{x!} = 1 - e^{-3} \left( 1 + 3 + \frac{3^2}{2!} \right) \\ &= .5768. \end{aligned}$$

- (b)  $\lambda = \frac{30}{7}$  per month.

$$P(X = 1) = e^{-\frac{30}{7}} \left( \frac{30}{7} \right) = .0590.$$

- (c) Albert is going to have at least 2 sales next month. Then we calculate the conditional probability

$$\begin{aligned} P(X \geq 4 / X \geq 2) &= \frac{P(X \geq 4 \ \& \ X \geq 2)}{P(X \geq 2)} = \frac{P(X \geq 4)}{P(X \geq 2)} = \frac{1 - \sum_{x=0}^3 P(X = x)}{1 - \sum_{x=0}^1 P(X = x)} \\ &= \frac{1 - e^{-\frac{30}{7}} \left[ 1 + \frac{30}{7} + \left( \frac{30}{7} \right)^2 \frac{1}{2!} + \left( \frac{30}{7} \right)^3 \frac{1}{3!} \right]}{1 - e^{-\frac{30}{7}} \left( 1 + \frac{30}{7} \right)} = .6689. \end{aligned}$$



### Example 2

Weak spots occur in a certain kind of weaved cloth on the average of one per 100 m. Assuming a Poisson distribution of the number of weak spots in any given length of cloth, what is the probability that

- a 240-m roll will have at most two defects?
- a 120-m roll will have no defects?
- Of five 120-m rolls, at least three of them will have no defects?
- Suppose a new weaving process can reduce weak spots so that any five 120-m rolls being free of defects has probability at least 0.1. Find the percentage reduction of weak spots.

### [Solution]

- (a) Let  $X_1$  be the number of defects per 240-m. Then its mean is 2.4.

$$\begin{aligned} P(X_1 \leq 2) &= P(X_1 = 0, 1, 2) \\ &= \sum_{x=0}^2 \frac{e^{-2.4} 2.4^x}{x!} = e^{-2.4} \left( 1 + 2.4 + \frac{2.4^2}{2!} \right) = .5697. \end{aligned}$$

- (b) Let  $X_2$  be the number of defects per 120-m. Then its mean is 1.2.

$$P(X_2 = 0) = \frac{e^{-1.2} 1.2^0}{0!} = .3012.$$

- (c) Let  $Y$  be the number of rolls with no defects out of five 120-m rolls. Then  $Y \sim b(5, 0.3012)$ , i.e.

$$P(Y = y) = C_y^5 (.3012)^y (1 - .3012)^{5-y}, \quad y = 0, 1, 2, \dots, 5.$$

$$\begin{aligned} &P(\text{at least three of them have no defects}) \\ &= P(Y \geq 3) = P(Y = 3) + P(Y = 4) + P(Y = 5) \\ &= C_3^5 (.3012)^3 (.6988)^2 + C_4^5 (.3012)^4 (.6988) + C_5^5 (.3012)^5 (.6988)^0 \\ &= .1334 + .0288 + .0025 = .1647. \end{aligned}$$

- (d) The mean becomes  $\lambda = 1.2(1-d)$  if  $100d\%$  weak spots are reduced.

Then  $P(X_2 = 0) = e^{-1.2(1-d)}$  by (b).

$$P(5 \text{ rolls free of defects}) = [e^{-1.2(1-d)}]^5 \geq .1,$$

$$\text{i.e. } e^{-6(1-d)} \geq .1.$$

On solving,  $-6(1-d) \geq \ln(.1)$

$$1-d \leq .3838, \quad d \geq .6162.$$

This means at least 61.62% weak spots are reduced.

### Example 3

An English teacher Miss Lee has been marking students' compositions for years. She finds that only 5% of the pages are free of errors (for example, grammatical and spelling errors, etc.). Suppose the number of errors per page follows a Poisson distribution.

- Show that on the average there are about 3 errors per page.
- What is the probability that a composition of 2 pages contains more than 4 errors?
- Another English teacher Miss Chan requires her students to hand in their compositions in typing. However, typing errors occurs. It is found that only 1% of the pages are free of errors. Assuming the number of typing errors per page follows a Poisson distribution.

Use relevant assumptions to calculate the mean number of typing errors per page.

### [Solution]

- Let  $X_1$  be the number of errors per page and  $\lambda$  be its mean.

$$P(X_1 = x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, 2, \dots$$

$$P(X_1 = 0) = e^{-\lambda} = .05$$

$$\lambda = -\ln(.05) = 2.9957 \approx 3.$$

- Let  $X_2$  be the number of errors per 2 pages. Its mean is 6.

$$\begin{aligned} P(X_2 > 4) &= 1 - \sum_{x=0}^4 P(X_2 = x) = 1 - e^{-6} \sum_{x=0}^4 \frac{6^x}{x!} \\ &= 1 - e^{-6} \left( 1 + 6 + \frac{6^2}{2!} + \frac{6^3}{3!} + \frac{6^4}{4!} \right) = .7149. \end{aligned}$$

- An error is either due to incorrect language usage (i.e. grammar, spelling, etc.) or mistyping. The case of both is not counted.
  - The content of one written page equals that of one typed page.

Let  $Y$  be the number of error per typed page and  $\mu$  be its mean.

$$P(Y = y) = \frac{e^{-\mu} \mu^y}{y!}, \quad P(Y = 0) = e^{-\mu} = .01 \Rightarrow \mu = 4.61.$$

The mean typing error per page is  $4.61 - 3 = 1.61$ .

**Example 4**

- (a) Between the hours of 2:30 p.m. and 5:00 p.m., the average number of calls per minute coming into a switchboard is 2.5. Assume the number of calls follow a Poisson distribution. Find the probabilities that during one particular minute there are
- (i) 4 or fewer calls, (ii) more than 6 calls.
- (b) During the period 3:00 p.m. to 3:02 p.m., show that the probability of receiving an odd number of telephone calls is approximately  $\frac{1}{2}$ .
- (c) Does it follow from (b) that the probability of receiving an even number of calls is also  $\frac{1}{2}$ ?
- (d) Suppose the operators answer, on the average,  $\frac{1}{5}$  of the incoming calls in English. What is the probability that they answer exactly two telephone calls in English in a given minute?

 **[Solution]**

- (a) Let  $X$  be the number of calls per minute.

$$P(X = x) = \frac{e^{-2.5} 2.5^x}{x!}, \quad x = 0, 1, 2, \dots$$

$$\begin{aligned} \text{(i) } P(X \leq 4) &= \sum_{x=0}^4 P(X = x) = \sum_{x=0}^4 \frac{e^{-2.5} 2.5^x}{x!} \\ &= e^{-2.5} \left( 1 + 2.5 + \frac{2.5^2}{2!} + \frac{2.5^3}{3!} + \frac{2.5^4}{4!} \right) = .8912. \end{aligned}$$

$$\begin{aligned} \text{(ii) } P(X > 6) &= 1 - P(X \leq 6) = 1 - \sum_{x=0}^6 \frac{e^{-2.5} 2.5^x}{x!} \\ &= 1 - e^{-2.5} \left( 1 + 2.5 + \frac{2.5^2}{2!} + \dots + \frac{2.5^6}{6!} \right) = .0142. \end{aligned}$$

- (b) During the 2-minute period, let  $Y$  be the number of telephone calls. It is a Poisson random variable with mean  $2.5 \times 2 = 5$  calls.

$$\begin{aligned} &\text{Probability of receiving an odd number of calls} \\ &= P(Y \text{ takes on an odd number}) \\ &= P(Y = 1, 3, 5, \dots) \\ &= P(Y = 1) + P(Y = 3) + P(Y = 5) + \dots \\ &= e^{-5} \left( 5 + \frac{5^3}{3!} + \frac{5^5}{5!} + \dots \right). \end{aligned}$$

Recall that

$$e^5 = 1 + 5 + \frac{5^2}{2!} + \frac{5^3}{3!} + \dots \dots \dots (1)$$

$$e^{-5} = 1 - 5 + \frac{5^2}{2!} - \frac{5^3}{3!} + \dots \dots \dots (2)$$

(1) - (2) gives

$$e^5 - e^{-5} = 2 \left( 5 + \frac{5^3}{3!} + \frac{5^5}{5!} + \dots \right).$$

Multiplying both sides by  $\frac{e^{-5}}{2}$ , we get

$$\frac{1 - e^{-10}}{2} = e^{-5} \left( 5 + \frac{5^3}{3!} + \frac{5^5}{5!} + \dots \right).$$

Note that LHS  $\approx \frac{1}{2}$  as  $e^{-10} \approx 0$ . (LHS = .4999773...)

Therefore RHS  $\approx \frac{1}{2}$ , which is the probability of receiving an odd number of calls.

(c) (1) + (2) gives

$$e^5 + e^{-5} = 2 \left( 1 + \frac{5^2}{2!} + \frac{5^4}{4!} + \dots \right).$$

Similarly to (b), we get

$$\frac{1 + e^{-10}}{2} = e^{-5} \left( 1 + \frac{5^2}{2!} + \frac{5^4}{4!} + \dots \right).$$

It follows that RHS  $\approx \frac{1}{2}$ , i.e.

$$P(X = 0) + P(X = 2) + P(X = 4) + \dots \approx \frac{1}{2}.$$

But 0 is not an even number. Hence,

$$\begin{aligned} P(X = 2) + P(X = 4) + \dots &\approx \frac{1}{2} - P(X = 0) \\ &= .5 - .0067. \\ &= .4933. \end{aligned}$$

The probability of receiving an even number of calls is slightly less than  $\frac{1}{2}$ .

(d) Let  $Y$  be the number of calls per minute “in English”.

If  $Y = 2$ , then  $X \geq 2$ . By the theorem of total probability on P. 68,

$$\begin{aligned}
 P(Y = 2) &= \sum_{x=2}^{\infty} P(X = x)P(Y = 2 / X = x) \\
 &= \sum_{x=2}^{\infty} \frac{e^{-2.5} 2.5^x}{x!} \cdot C_2^x \left(\frac{1}{5}\right)^2 \left(\frac{4}{5}\right)^{x-2} \left[ Y \sim b\left(x, \frac{1}{5}\right) \text{ when } X = x \right] \\
 &= \sum_{x=2}^{\infty} \frac{e^{-2.5} 2.5^x}{x!} \cdot \frac{x!}{2!(x-2)!} \left(\frac{1}{5}\right)^2 \left(\frac{4}{5}\right)^{x-2} \\
 &= \sum_{x=2}^{\infty} \frac{e^{-2.5} 2.5^2}{2!(x-2)!} \cdot \frac{1}{5^2} \left(2.5 \times \frac{4}{5}\right)^{x-2} \\
 &= \sum_{x=2}^{\infty} \frac{e^{-2.5}}{2} \left(\frac{2.5}{5}\right)^2 \cdot \frac{2^{x-2}}{(x-2)!} \\
 &= \frac{e^{-2.5}}{2 \cdot 4} \sum_{x=2}^{\infty} \frac{2^{x-2}}{(x-2)!} \\
 &= \frac{e^{-2.5}}{2 \cdot 4} \left(1 + 2 + \frac{2^2}{2!} + \frac{2^3}{3!} + \dots\right) \\
 &= \frac{e^{-2.5}}{2 \cdot 4} \cdot e^2 = \frac{e^{-.5}}{8} \\
 &= .0758.
 \end{aligned}$$

**Note:**

We can go further to prove that

$$P(Y = 2) = \sum_{x=y}^{\infty} P(X = x)P(Y = y / X = x) = \frac{e^{-.5} (.5)^y}{y!}, \quad y = 0, 1, 2, \dots,$$

i.e.  $Y$  has a Poisson distribution with mean 0.5.

Steps of the proof parallel those of above.

$$\begin{aligned}
 \sum_{x=y}^{\infty} P(X = x)P(Y = y / X = x) &= \sum_{x=y}^{\infty} \frac{e^{-2.5} 2.5^x}{x!} C_y^x \left(\frac{1}{5}\right)^y \left(\frac{4}{5}\right)^{x-y} \\
 &= \sum_{x=y}^{\infty} \frac{e^{-2.5} 2.5^x}{x!} \cdot \frac{x!}{y!(x-y)!} \left(\frac{1}{5}\right)^y \left(\frac{4}{5}\right)^{x-y} \\
 &\quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots
 \end{aligned}$$

Further reduction is left to the student.

## IV. THE GEOMETRIC DISTRIBUTION

---

Suppose independent Bernoulli trials with probability of success  $p$  are conducted. Let  $X$  be the number of such Bernoulli trials until we get a success. Then  $X$  is said to have a geometric distribution with parameter  $p$ . Observe that there are  $X - 1$  failures preceding a success. Thus the probability function of  $X$  is

$$P(X = x) = (1 - p)^{x-1} p, \quad x = 1, 2, \dots$$

Its mean is  $\frac{1}{p}$  and variance is  $\frac{1}{p} \left( \frac{1}{p} - 1 \right)$ .

### Example 1

An American roulette wheel commonly has 38 spots on it of which 18 are black, 18 are red and 2 are green. Let  $X$  and  $Y$  be the number of spins necessary to observe the first red and first green number respectively. Find the probability functions, means and variances for  $X$  and  $Y$ .

### [Solution]

The probability of getting a red number is  $\frac{18}{38}$ , while that of getting a green number is  $\frac{2}{38}$ .

Then we have

$$(i) \quad P(X = x) = \left( \frac{20}{38} \right)^{x-1} \left( \frac{18}{38} \right) = \left( \frac{10}{19} \right)^{x-1} \left( \frac{9}{19} \right), \quad x = 1, 2, \dots$$

$$\mu_x = \frac{38}{18} = \frac{19}{9}, \quad \sigma_x^2 = \frac{38}{18} \left( \frac{38}{18} - 1 \right) = \frac{38}{18} \cdot \frac{20}{18} = \frac{190}{81}.$$

$$(ii) \quad P(Y = y) = \left( \frac{36}{38} \right)^{y-1} \left( \frac{2}{38} \right) = \left( \frac{18}{19} \right)^{y-1} \left( \frac{1}{19} \right), \quad y = 1, 2, \dots$$

$$\mu_y = \frac{38}{2} = 19, \quad \sigma_y^2 = 19(19 - 1) = 342.$$

**Example 2**

A fair die is rolled until a 6 occurs. Compute the probability that

- 10 rolls are needed.
- less than 4 rolls are needed.
- an odd number of rolls is needed.

 **[Solution]**

Let  $X$  be the number of rolls needed.

$$P(X = x) = \left(\frac{5}{6}\right)^{x-1} \frac{1}{6}, \quad x = 1, 2, \dots$$

$$(a) \quad P(X = 10) = \left(\frac{5}{6}\right)^9 \frac{1}{6} = .0323.$$

$$(b) \quad P(X < 4) = P(X = 1) + P(X = 2) + P(X = 3)$$

$$= \frac{1}{6} + \frac{5}{6} \cdot \frac{1}{6} + \left(\frac{5}{6}\right)^2 \frac{1}{6} = .4213.$$

$$(c) \quad P(X \text{ is an odd number}) = P(X = 1) + P(X = 3) + P(X = 5) + \dots$$

$$= \frac{1}{6} + \left(\frac{5}{6}\right)^2 \frac{1}{6} + \left(\frac{5}{6}\right)^4 \frac{1}{6} + \dots$$

$$= \frac{1}{6} \left[ 1 + \left(\frac{5}{6}\right)^2 + \left(\frac{5}{6}\right)^4 + \dots \right]$$

$$= \frac{1}{6} \cdot \frac{1}{1 - \left(\frac{5}{6}\right)^2} = \frac{1}{6} \cdot \frac{36}{36 - 25} = \frac{6}{11}.$$

未完待續