Chapter 5
Some Discrete Probability Distributions

吳明龍 助理教授
成大資訊系/醫資所
minglong.wu@csie.ncku.edu.tw
辦公室:雲平大樓東棟 415 室

5.4 Negative Binomial and Geometric Distributions

- Negative binomial experiments: the *k*th success occurs on the *x*th trial.
- Negative binomial random variable: the number *X* of trials to produce *k* success in a negative binomial experiment.
- Negative binomial distribution: If repeated independent trials can result in a success with probability *p* and a failure with probability *q* = 1 – *p*, then the probability distribution of the random variable *X*, the number of the trial on which the *k*th success occurs, is

$$b^*(x;k,p) = \binom{x-1}{k-1} p^k q^{x-k}, \quad x = k, k+1, k+2, \cdots$$

Negative Binomial and Geometric Distributions

- Example 5.14: In an NBA (National Basketball Association) championship series, the team which wins four games out of seven will be the winner. Suppose that team *A* has probability 0.55 of winning over the team B and both teams *A* and *B* face each other in championship games.
  (a) What is the probability that team A will win the series in six games?
  (b) What is the probability that team A will win the series?
  (c) If both teams face each other in a regional playoff series and the winner is decided by winning three out of five games, what is the probability that team *A* will win a playoff?

(a) $b^*(6;4,0.55) = \binom{5}{3} 0.55^4 (1-0.55)^{6-4} = 0.1853$

(b) $b^*(4;4,0.55) + b^*(5;4,0.55) + b^*(6;4,0.55) + b^*(7;4,0.55)$
$= 0.0915 + 0.1647 + 0.1853 + 0.1668 = 0.6083$

(c) $b^*(3;3,0.55) + b^*(4;3,0.55) + b^*(5;3,0.55)$
$= 0.1664 + 0.2246 + 0.2021 = 0.5931$

Negative Binomial and Geometric Distributions

- Geometric Distribution: If repeated independent trials can result in a success with probability *p* and a failure with probability *q* = 1 – *p*, then the probability distribution of the random variable X, the number of the trial on which the first success occurs, is

$$g(x;p) = b^*(x;1,p) = pq^{x-1}, \quad x = 1,2,3,\cdots$$

- Example 5.15: In a certain manufacturing process it is known that, on the average, 1 in every 100 items is defective. What is the probability that the fifth item inspected is the first defective item found?

$$g(5;0.01) = 0.01 \cdot 0.99^4 = 0.0096$$

Negative Binomial and Geometric Distributions

- Example 5.16: At “busy time” a telephone exchange is very near capacity, so callers have difficulty placing their calls. It may be of interest to know the number of attempts necessary in order to gain a connection. Suppose that let p = 0.05 be the probability of a connection during busy time. We are interested in knowing the probability that 5 attempts are necessary for a successful call.
  - **Solution** $P(X = x) = g(5;0.05) = 0.05 \cdot 0.95^4 = 0.041$

- Theorem 5.3: The mean and variance of a random variable following the geometric distribution are

$$\mu = \frac{1}{p}, \qquad \sigma^2 = \frac{1-p}{p^2}$$

- In the system of telephone exchange, trials occurring prior to a success represent a cost.
- A high probability of requiring a large of number of attempts is not beneficial to the scientists or engineers.

5.5 Poisson Distribution and the Poisson Process

- Poisson experiments: Experiments yielding numerical values of a random variable X, the number of outcomes occurring <u>during a given time interval or in a specified region</u>.
- Properties of Poisson Process:

1. The number of outcomes in one time interval or specified region is <u>independent</u> of the number that occurs in any other disjoint time interval or region of space. In this way we say that the Poisson process has <u>no memory</u>.
2. The probability that a single outcome will occur during a very short time interval or in a small region <u>is proportional to the length of the time interval or the size of the region</u> and does not depend on the number of outcomes occurring outside this time interval or region.
3. The probability that more than one outcome will occur in such a short time interval or fall in such a small region is negligible.

Poisson Distribution and the Poisson Process

- Poisson Distribution: The probability distribution of the Poisson random variable *X*, representing the number of outcomes occurring in a given time interval or specified region denoted by *t*, is (mean number $\mu = \lambda t$)

$$p(x;\lambda t) = \boxed{\frac{e^{-\lambda t}(\lambda t)^x}{x!}}, \quad x = 0,1,2,\cdots$$

where λ is the average number of outcomes per unit time or region, and $e = 2.71828$.

Poisson Distribution and the Poisson Process

- Example 5.17: During a laboratory experiment the average number of radioactive particles passing through a counter in 1 millisecond is 4. What is the probability that 6 particles enter the counter in a given millisecond? (Table A.2)

$$p(6;4) = \frac{e^{-4}4^6}{6!} = \sum_{x=0}^{6} p(x;4) - \sum_{x=0}^{5} p(x;4) = 0.8893 - 0.7851 = 0.1042$$

(p.732)

- Example 5.18: <u>Ten</u> is the average number of oil tankers arriving each day at a certain port city. The facilities at the port can handle at most 15 tankers per day. What is the probability that on a given day tankers have to be turned away?

$$P(X > 15) = 1 - P(X \le 15) = 1 - \sum_{x=0}^{15} p(x;10) = 1 - 0.9513 = 0.0487$$

(p.734)

TABLE A.2 Poisson Probability Sums $\sum_{x=0}^{r} p(x; \mu)$

	μ								
r	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0	0.9048	0.8187	0.7408	0.6730	0.6065	0.5488	0.4966	0.4493	0.4066
1	0.9953	0.9825	0.9631	0.9384	0.9098	0.8781	0.8442	0.8088	0.7725
2	0.9998	0.9989	0.9964	0.9921	0.9856	0.9769	0.9659	0.9526	0.9371
3	1.0000	0.9999	0.9997	0.9992	0.9982	0.9966	0.9942	0.9909	0.9865
4		1.0000	1.0000	0.9999	0.9998	0.9996	0.9992	0.9986	0.9977
5				1.0000	1.0000	1.0000	0.9999	0.9998	0.9997
6							1.0000	1.0000	1.0000

	μ								
r	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
0	0.3679	0.2231	0.1353	0.0821	0.0498	0.0302	0.0183	0.0111	0.0067
1	0.7358	0.5578	0.4060	0.2873	0.1991	0.1359	0.0916	0.0611	0.0404
2	0.9197	0.8088	0.6767	0.5438	0.4232	0.3208	0.2381	0.1736	0.1247
3	0.9810	0.9344	0.8571	0.7576	0.6472	0.5366	0.4335	0.3423	0.2650
4	0.9963	0.9814	0.9473	0.8912	0.8153	0.7254	0.6288	0.5321	0.4405
5	0.9994	0.9955	0.9834	0.9580	0.9161	0.8576	0.7851	0.7029	0.6160
6	0.9999	0.9991	0.9955	0.9858	0.9665	0.9347	0.8893	0.8311	0.7622
7	1.0000	0.9998	0.9989	0.9958	0.9881	0.9733	0.9489	0.9134	0.8666
8		1.0000	0.9998	0.9989	0.9962	0.9901	0.9786	0.9597	0.9319
9			1.0000	0.9997	0.9989	0.9967	0.9919	0.9829	0.9682
10				0.9999	0.9997	0.9990	0.9972	0.9933	0.9863
11				1.0000	0.9999	0.9997	0.9991	0.9976	0.9945
12					1.0000	0.9999	0.9997	0.9992	0.9980
13						1.0000	0.9999	0.9997	0.9993
14							1.0000	0.9999	0.9998
15								1.0000	0.9999
16									1.0000

Poisson Distribution and the Poisson Process

- Theorem 5.4: The mean and variance of the Poisson distribution

  $p(x;\lambda t)$ both have the value λt.

  – Proof in Appendix A.25 (p.768)

  $\mu = \sigma^2 = \lambda t$

- Example :

  In Example 5.20, $\lambda t = 4 \Rightarrow \mu = \sigma^2 = 4, \sigma = 2 \Rightarrow \mu \pm 2\sigma = 4 \pm 2 \cdot 2$ (from 0 to 8)

  Using Chebyshev's theorem, we conclude that at least 3/4 of the time the number of radioactive particles entering the counter will be anywhere from 0 to 8 during a given millisecond.

A.25

PROOF To verify that the mean is indeed λt, let $\mu = \lambda t$. We can write

$$E(X) = \sum_{x=0}^{\infty} x \cdot \frac{e^{-\mu}\mu^x}{x!} = \sum_{x=1}^{\infty} x \cdot \frac{e^{-\mu}\mu^x}{x!} = \mu \sum_{x=1}^{\infty} \frac{e^{-\mu}\mu^{x-1}}{(x-1)!}.$$

Now, let $y = x - 1$ to give

$$E(X) = \mu \sum_{y=0}^{\infty} \frac{e^{-\mu}\mu^y}{y!} = \mu,$$

since

$$\sum_{y=0}^{\infty} \frac{e^{-\mu}\mu^y}{y!} = \sum_{y=0}^{\infty} p(y;\mu) = 1.$$

The variance of the Poisson distribution is obtained by first finding

$$\underline{E[X(X-1)]} = \sum_{x=0}^{\infty} x(x-1)\frac{e^{-\mu}\mu^x}{x!} = \sum_{x=2}^{\infty} x(x-1)\frac{e^{-\mu}\mu^x}{x!}$$

$$= \mu^2 \sum_{x=2}^{\infty} \frac{e^{-\mu}\mu^{x-2}}{(x-2)!}.$$

Setting $y = x - 2$, we have

$$E[X(X-1)] = \mu^2 \sum_{y=0}^{\infty} \frac{e^{-\mu}\mu^y}{y!} = \mu^2.$$

Hence

$$\sigma^2 = E[X(X-1)] + \mu - \mu^2 = \mu^2 + \mu - \mu^2 = \mu = \lambda t.$$

Nature of the Poisson Probability Function

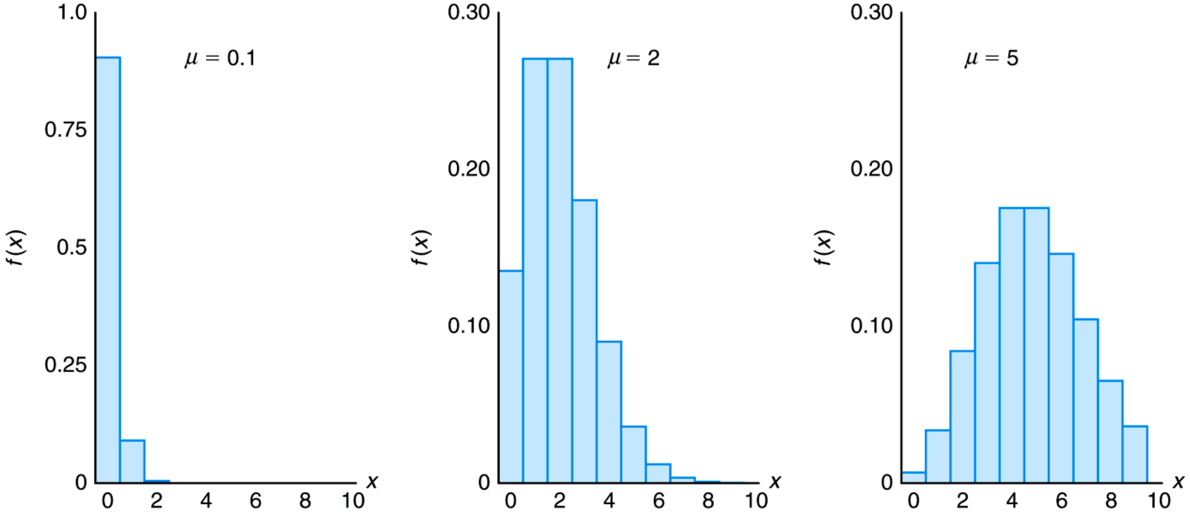

Figure 5.1: Poisson density functions for different means.

Approximation of Binomial Distribution by a Poisson Distribution

- Theorem 5.5: Let *X* be a binomial random variable with probability distribution *b*(*x*; *n*, *p*).

$$\text{When } n \to \infty, p \to 0, \text{and } \mu = np \text{ remains constant.}$$

$$b(x; n, p) \to p(x; \mu).$$

  – Proof is in Appendix A27 (8^{th} ed.).

- Example 5.19: In a certain industrial facility accidents occur infrequently. It is known that the probability of an accident on any given day is 0.005 and accidents are independent of each other.
  (a) What is the probability that in any given period of 400 days there will be an accident on one day?
  (b) What is the probability that there are at most three days with an accident?

  – **Solution**

$$n = 400, p = 0.005 \Rightarrow np = 2, \text{using Poisson approximation}$$

$$\text{(a) } P(X = 1) = e^{-2} 2^1 = 0.271$$

$$\text{(b) } P(X \le 3) = \sum_{x=0}^{3} \frac{e^{-2} 2^x}{x!} = 0.857$$

(A.27)

PROOF The binomial distribution can be written

$$b(x; n, p) = \binom{n}{x} p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

$$= \frac{n(n-1)\cdots(n-x+1)}{x!} p^x (1-p)^{n-x}.$$

Substituting $p = \mu/n$, we have

$$b(x; n, p) = \frac{n(n-1)\cdots(n-x+1)}{x!} \left(\frac{\mu}{n}\right)^x \left(1 - \frac{\mu}{n}\right)^{n-x}$$

$$= 1\left(1 - \frac{1}{n}\right)\cdots\left(1 - \frac{x-1}{n}\right) \boxed{\frac{\mu^x}{x!}} \left(1 - \frac{\mu}{n}\right)^n \left(1 - \frac{\mu}{n}\right)^{-x}.$$

As $n \to \infty$ while x and μ remain constant,

$$\lim_{n\to\infty} 1\left(1 - \frac{1}{n}\right)\cdots\left(1 - \frac{x-1}{n}\right) = 1, \quad \lim_{n\to\infty} \left(1 - \frac{\mu}{n}\right)^{-x} = 1,$$

and from the definition of the number e,

$$(1 + \frac{x}{n})^n \to e^x \text{ as } n \to \infty$$

$$\lim_{n\to\infty} \left(1 - \frac{\mu}{n}\right)^n = \lim_{n\to\infty} \left\{\left[1 + \frac{1}{(-n)/\mu}\right]^{-n/\mu}\right\}^{-\mu} = e^{-\mu}.$$

Hence, under the given limiting conditions,

$$b(x; n, p) \to \frac{e^{-\mu}\mu^x}{x!}, \qquad x = 0, 1, 2, \ldots$$

Approximation of Binomial Distribution by a Poisson Distribution

- Example 5.20: In a manufacturing process where glass products are produced, defects or bubbles occur, occasionally rendering the piece undesirable for marketing. It is known that, on average, 1 in every 1000 of these items produced has one or more bubbles. What is the probability that a random sample of 8000 will yield fewer than 7 items processing bubbles?
  - **Solution**

$$n = 8000,\ p = 0.001 \Rightarrow \mu = np = 8, \text{ using Poisson approximation}$$

$$P(X < 7) = \textstyle\sum_{x=0}^{6} b(x;8000,0.001) \approx \sum_{x=0}^{6} p(x;8) = 0.3134.$$

Exercise

- 5.49, 5.55, 5.65(by both table and Matlab)