Chapter 3

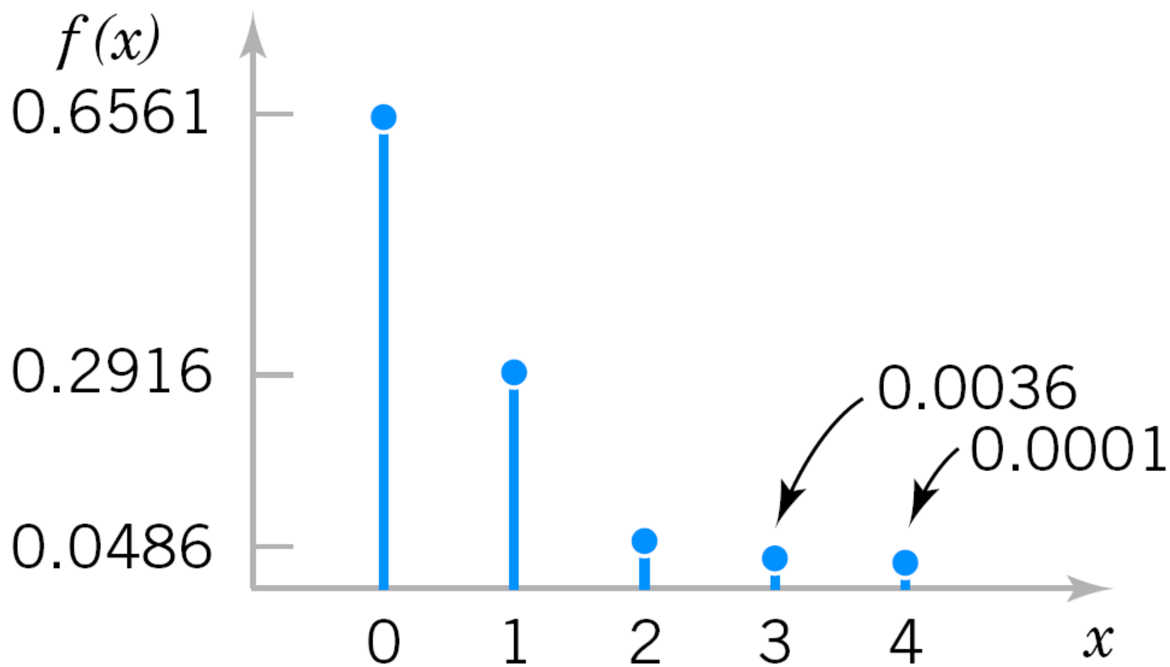
- 3-2 Probability distributions and probability mass functions
- 3-3 Cumulative distribution functions
- 3-4 Mean and variance of a discrete random variable
- 3-5 Discrete Uniform Distribution
- 3-6 Binomial Distribution
- 3-7 Geometric and Negative Binomial Distribution
- 3-8 Hypergeometric Distribution
- 3-9 Poisson Distribution

3-2 probability distributions and probability mass functions

- Example 3-4

There is a chance that a bit transmitted through a digital transmission channel is received in error. Let X equal the number of bits in error in the next four bits transmitted. The possible values for X are $\{0, 1, 2, 3, 4\}$.

- $P(X=0)=0.6561$,
  $P(X=1)=0.2916$,
  $P(X=2)=0.0486$,
  $P(X=3)=0.0036$,
  $P(X=4)=0.0001$



- Definition

For a discrete random variable X with possible values $x_1, x_2, \ldots, x_n$, a **probability mass function** is a function such that

$$(1) \quad f(x_i) \geq 0$$

$$(2) \quad \sum_{i=1}^{n} f(x_i) = 1$$

$$(3) \quad f(x_i) = P(X = x_i) \qquad (3\text{-}1)$$

➢ In example 3-4, $f(0)=0.6561$, $f(1)=0.2916$, $f(2)=0.0486$, $f(3)=0.0036$, $f(4)=0.0001$.

3-3 Cumulative distribution functions

- Ex. 3-6 $P(X \leq 3)$ for Ex.3-4
  - The event that $\{X \leq 3\}$ is the union of the events $\{X=0\}$, $\{X=1\}$, $\{X=2\}$, and $\{X=3\}$.
  - $P(X \leq 3) = P(X=0) + P(X=1) + P(X=2) + P(X=3) = 0.9999$
- Definition

The **cumulative distribution function** of a discrete random variable X, denoted as $F(x)$, is

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$$

For a discrete random variable X, $F(x)$ satisfies the following properties.

(1) $F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$

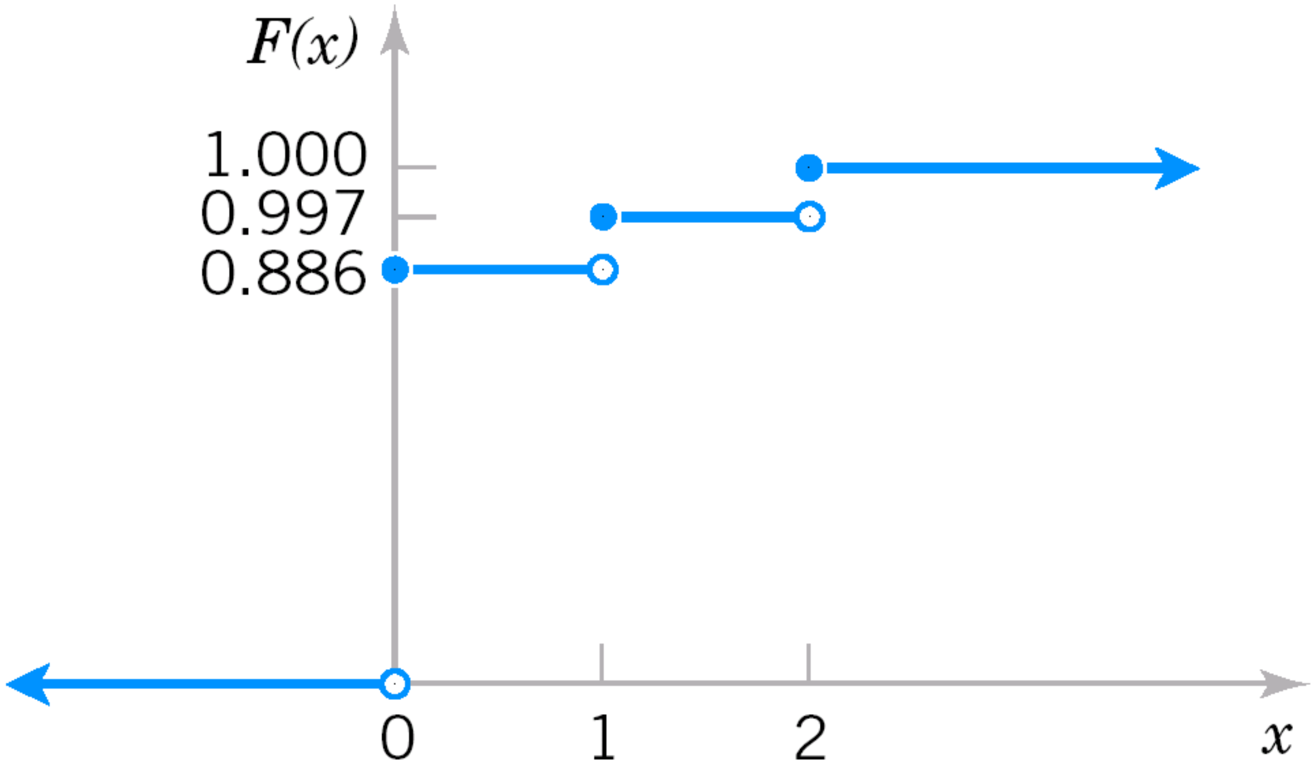
(2) $0 \leq F(x) \leq 1$

(3) If $x \leq y$, then $F(x) \leq F(y)$ (3-2)

Ex 3-8 Suppose that a day's production of 850 manufactured parts contains 50 parts that do not conform to customer requirements. Two parts are selected at random, without replacement, from the batch. Let the random variable X equal the number of nonconforming parts in the sample. What is the cumulative distribution function of X?

- $P(X=0) = (800/850)(799/849) = 0.886$
- $P(X=1) = (800/850)(50/849) + (50/850)(800/849) = 0.111$
- $P(X=2) = (50/850)(49/849) = 0.003$
- $F(0) = P(X=0) = 0.886$
- $F(1) = P(X \le 1) = 0.886 + 0.111 = 0.997$
- $F(2) = P(X \le 2) = 1$

- Note that $F(x)$ is defined for all x from $-\infty<x<\infty$ and not only for 0,1, and 2.



3-4 mean and variance of a discrete random variable

- Definition

The **mean** or **expected value** of the discrete random variable X, denoted as μ or $E(X)$, is

$$\mu = E(X) = \sum_x x f(x) \tag{3-3}$$

The **variance** of X, denoted as σ^2 or $V(X)$, is

$$\sigma^2 = V(X) = E[(X\text{-}\mu)^2] = \sum_x (x-\mu)^2 f(x) = \sum_x x^2 f(x) - \mu^2$$

The **standard deviation** of X is $\sigma = \sqrt{\sigma^2}$.

- ➢ The mean is a measure of the center or middle of the probability distribution.
- ➢ The variance is a measure of the dispersion, or variability in the distribution.

- Ex.3-9 for Ex.3-4
  - $\mu=E[X]=0f(0)+1f(1)+2f(2)+3f(3)+4f(4)=0.4$
  - $V(X)=\sigma^2=\sum_{i=1}^{5} f(x_i)(x_i-0.4)^2=0.36$
  - The variance of a random variable *X* can be considered to be the expected value of a specific function of *X*, namely, $h(X)=(X-\mu)^2$. In general, the expected value of any function $h(X)$ of a discrete random variable is defined in a similar manner.
  - Expected value of a function of a discrete random variable
    - ☞If *X* is a discrete random variable with probability mass function $f(x)$

$$E[h(x)]=\sum_x h(x)f(x)$$

3-5 Discrete Uniform Distribution

- Definition

A random variable X has a **discrete uniform distribution** if each of the n values in its range, say, $x_1, x_2, \ldots, x_n$, has equal probability. Then,

$$f(x_i) = 1/n \tag{3-5}$$

- Suppose the range of the discrete random variable X is the consecutive integers a, a+1, a+2,…, b for $a \leq b$. The range of X contains b-a+1 values each with probability 1/(b-a+1).

Suppose X is a discrete uniform random variable on the consecutive integers $a, a + 1, a + 2, \ldots, b$, for $a \leq b$. The mean of X is

$$\mu = E(X) = \frac{b + a}{2}$$

The variance of X is

$$\sigma^2 = \frac{(b - a + 1)^2 - 1}{12}$$

$$\sum_{n=1}^{N} n = \frac{N(N+1)}{2}, \quad \sum_{n=1}^{N} n^2 = \frac{N(N+1)(2N+1)}{6}$$

$$\sum_{n=a}^{b} n^2 = \sum_{n=0}^{b-a} (a+n)^2 = \sum_{n=0}^{b-a} a^2 + 2an + n^2 = (b-a+1)a^2 + 2a\frac{(b-a)(b-a+1)}{2}$$

$$+ \frac{(b-a)(b-a+1)(2b-2a+1)}{6}$$

$$= (b-a+1)[2a^2 + 2ab + 2b^2 + b - a]/6$$

3-6 Binomial Distribution

- Ex. Flip a coin 10 times. Let X=number of heads obtained.
- Ex. Of all bits transmitted through a digital transmission channel, 10% are received in error. Let X=the number of bits in error in the next five bits transmitted.
- **Bernoulli trial**
  - A trial with only two possible outcomes is used so frequently as a building block of a random experiment that it is called a **Bernoulli trial.**
  - It is usually assumed that the trials that constitute the random experiment are **independent.**
  - It is often reasonable to assume that the **probability of a success in each trial is constant.**

- Ex.3-16 The chance that a bit transmitted through a digital transmission channel is received in error is 0.1. Also, assume that the transmission trials are independent. Let X=the number of bits in error in the next four bits transmitted. Determine $P(X=2)$.
  - Let the letter E denote a bit in error, and let the letter O denote that the bit is okay, that is, received without error.
  - The event that X=2 consists of the six outcomes: $\{EEOO, EOEO, EOOE, OEEO, OEOE, OOEE\}$
  - $P(EEOO)=P(E)P(E)P(O)P(O)=(0.1)^2(0.9)^2=0.0081$
  - $P(X=2)=6(0.0081)=0.0486$
  - In general,

    $P(X=x)$=(number of outcomes that result in x errors) times $(0.1)^x(0.9)^{4-x}$.

- An outcome that contains *x* errors can be constructed by partitioning the four trials (letters) in the outcome into two groups. One group is of size *x* and contains the errors, and the other group is of size *n*-*x* and consists of the trials that are okay.
- The number of ways of partitioning four objects into two groups, one of which is of size *x*, is

$$\binom{4}{x} = \frac{4!}{x!(4-x)!} \quad \text{and, } P(X = x) = \binom{4}{x} 0.1^x 0.9^{4-x}$$

- $$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

- # Definition

A random experiment consists of n Bernoulli trials such that

(1) The trials are independent

(2) Each trial results in only two possible outcomes, labeled as "success" and "failure"

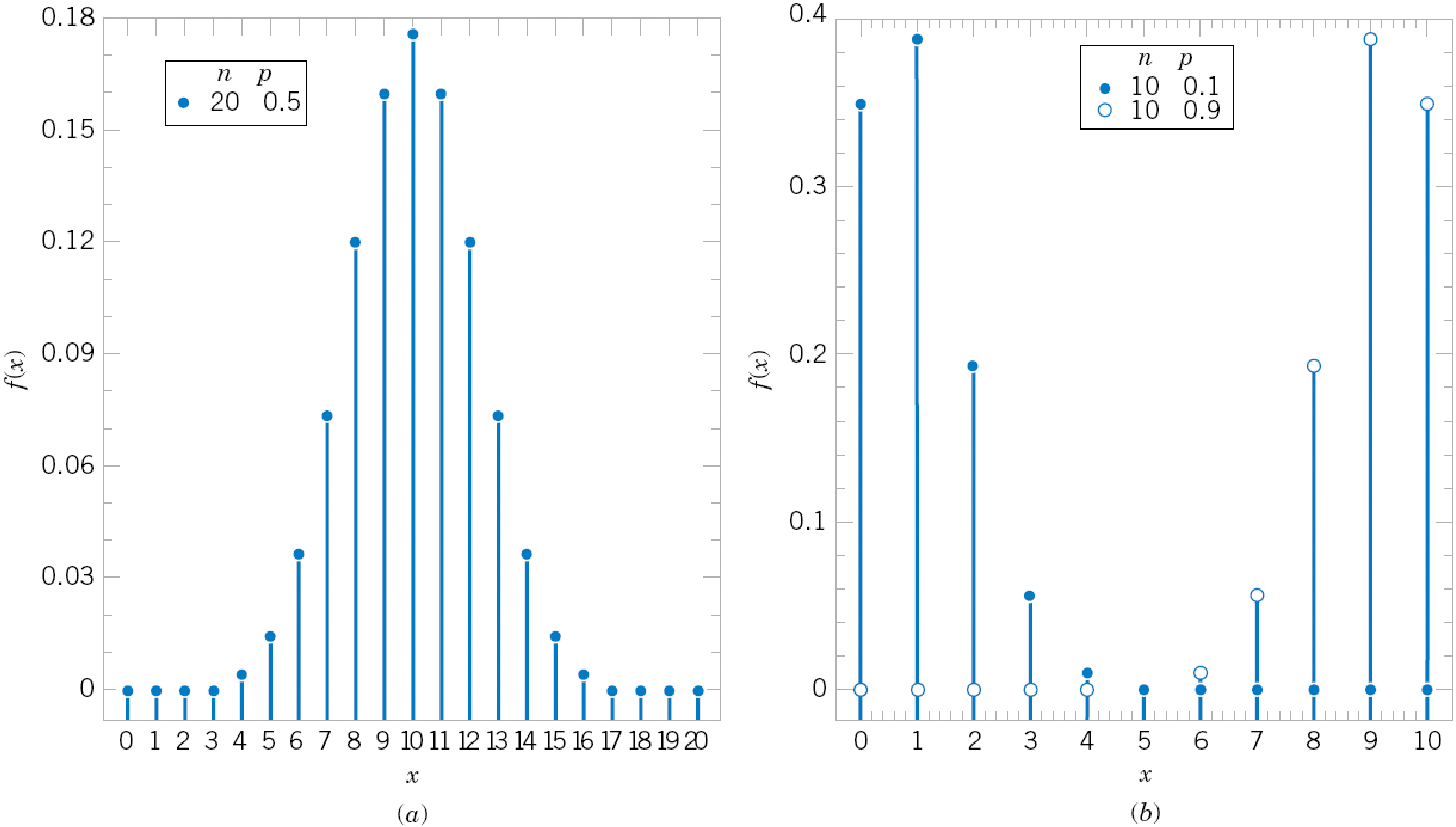
(3) The probability of a success in each trial, denoted as p, remains constant

The random variable X that equals the number of trials that result in a success has a **binomial random variable** with parameters $0 < p < 1$ and $n = 1, 2, \ldots$. The probability mass function of X is

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad x = 0, 1, \ldots, n \tag{3-7}$$

➢ If X is a binomial random variable with parameters p and n,

$$\mu = E(X) = np \qquad \text{and} \qquad \sigma^2 = V(X) = np(1-p)$$


0.18
0.15
0.12
0.09
0.06
0.03
0
f(x)
n p
20 0.5
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
x
(a)
0.4
0.3
0.2
0.1
0
f(x)
n p
10 0.1
10 0.9
0 1 2 3 4 5 6 7 8 9 10
x
(b)

- Ex. 3-18. Each sample of water has a 10% chance of containing a particular organic pollutant. Assume that the samples are independent with regard to the presence of the pollutant. Find the probability that in the next 18 samples,

(a) exactly 2 contain the pollutant.

(b) Determine the probability that at least four samples contain the pollutant.

- (a) $P(X=2)=\binom{18}{2}(0.1)^2(0.9)^{16}=\frac{18!}{(2!)(16!)}(0.1)^2(0.9)^{16}=0.284$
- (b) $P(X\geq 4)=\sum_{x=4}^{18}\binom{18}{x}(0.1^x)(0.9)^{18-x}$ = 1-$P(X<4)$

$$=1-\sum_{x=0}^{3}\binom{18}{x}(0.1)^x(0.9)^{18-x}=0.098$$

3-7 Geometric and Negative Binomial Distribution

- **Geometric distribution**
  - Instead of a fixed number of trials, trials are conducted until a success is obtained.

- Ex. 3-20 The probability that a bit transmitted through a digital transmission channel is received in error is 0.1. Assume the transmissions are independent events, and let the random variable *X* denote the number of bits transmitted *until* the first error.

  - $P(X=5)$ is the probability that the first four bits are transmitted correctly and the fifth bit is in error. $\{OOOOE\}$

  - $P(X=5)=0.9^4 0.1$

Note that there is some probability that X will equal any integer value. Also, if the first trial is a success, X=1. Therefore, the range of X is {1,2,3…}, that is, all positive integers.

- Definition

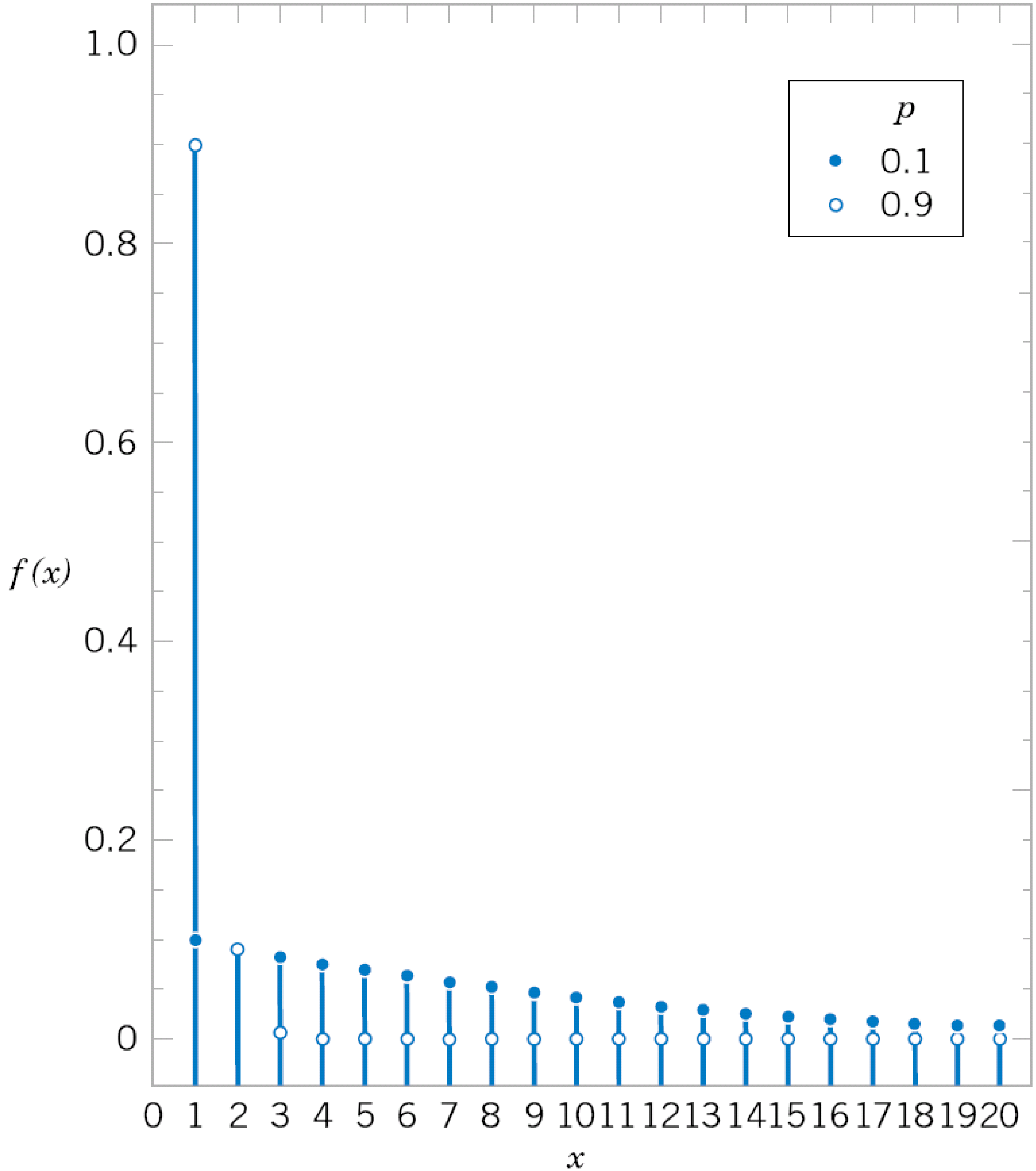
In a series of Bernoulli trials (independent trials with constant probability p of a success), let the random variable X denote the number of trials until the first success. Then X is a **geometric random variable** with parameter $0 < p < 1$ and

$$f(x) = (1-p)^{x-1}p \qquad x = 1, 2, \ldots \tag{3-9}$$

If X is a geometric random variable with parameter p,

$$\mu = E(X) = 1/p \qquad \text{and} \qquad \sigma^2 = V(X) = (1-p)/p^2$$

- Lack of Memory Property
  - ☞The probability of an error remains constant for all transmission. In this sense, the geometric distribution is said to lack any memory.



1.0
0.8
0.6
0.4
0.2
0
f (x)
p
0.1
0.9
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
x

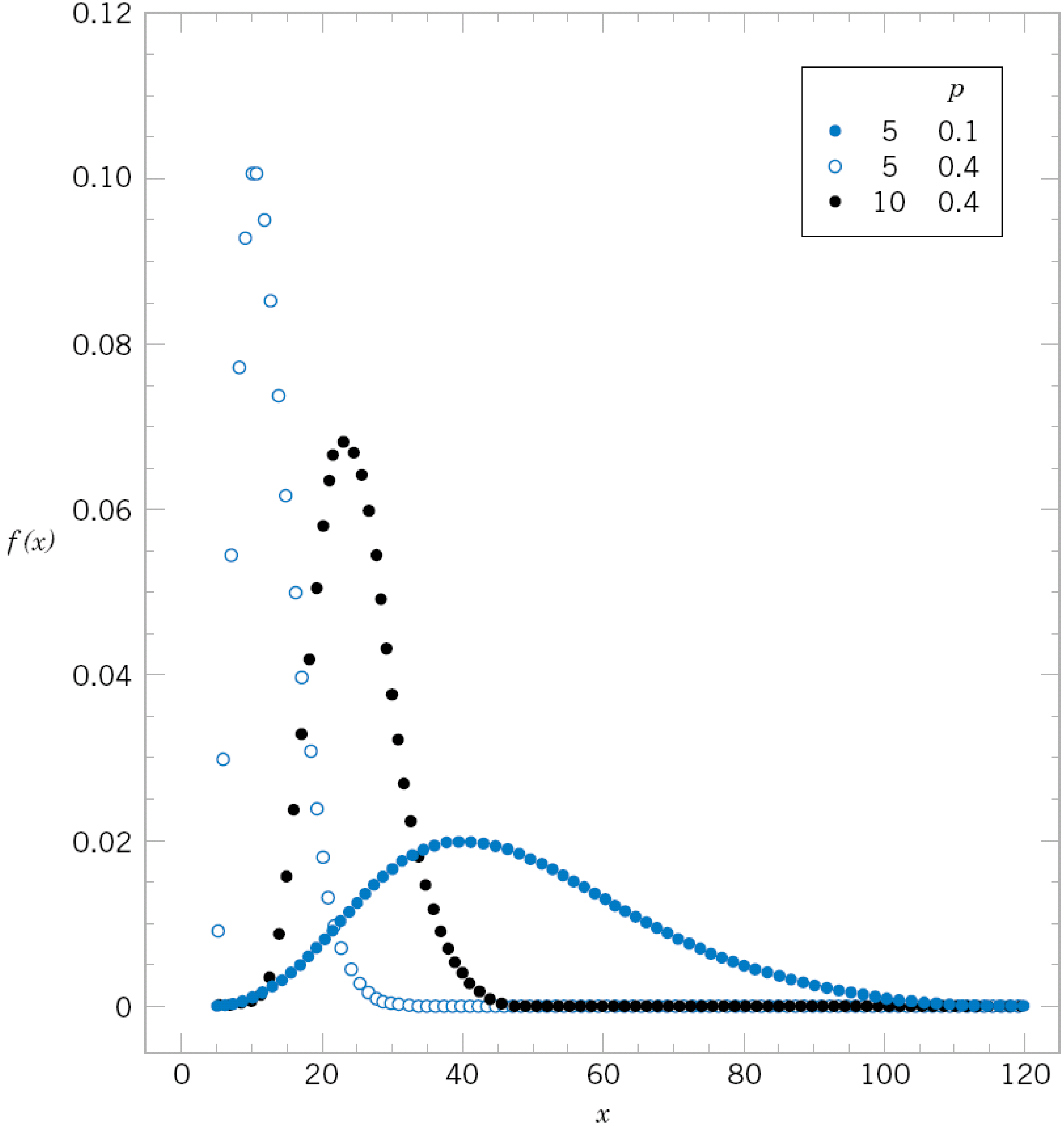
- # Negative Binomial Distribution

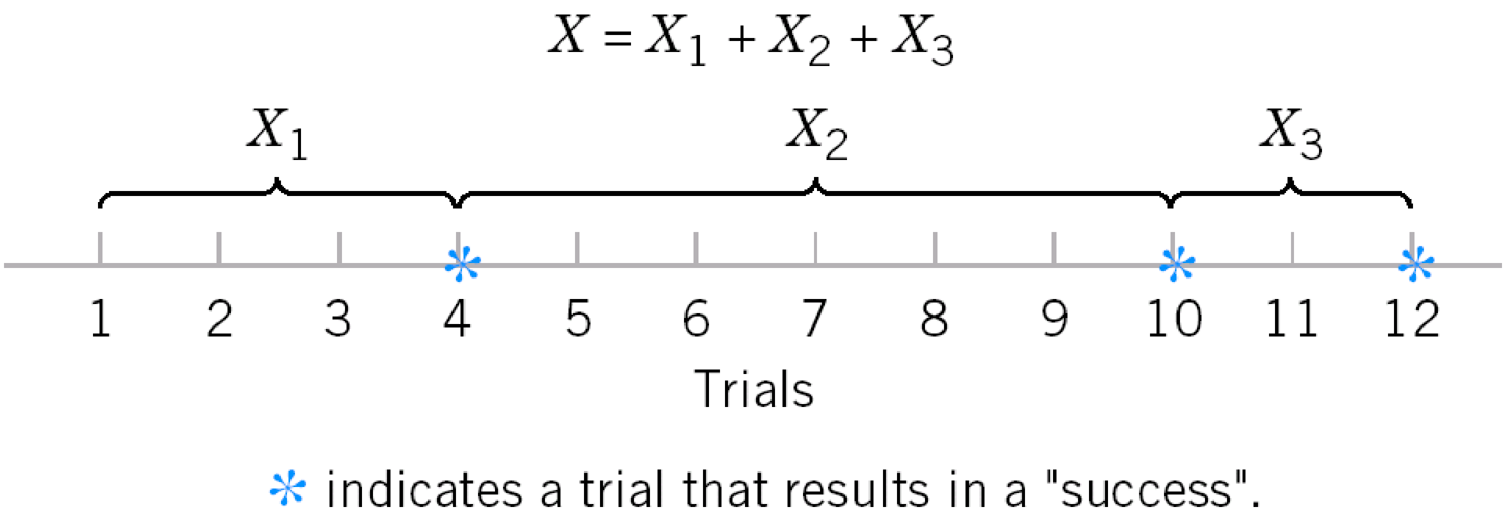
Definition

In a series of Bernoulli trials (independent trials with constant probability p of a success), let the random variable X denote the number of trials until r successes occur. Then X is a **negative binomial random variable** with parameters $0 < p < 1$ and $r = 1, 2\ 3, \ldots,$ and

$$f(x) = \binom{x-1}{r-1}(1-p)^{x-r}p^r \qquad x = r, r+1, r+2, \ldots. \tag{3-11}$$

- In the special case that r=1, a negative binomial random variable is a geometric random variable.
- A negative binomial random variable can be interpreted as the sum of r geometric random variables.
  - ☞The lack of memory property of a geometric random variable. (P.90)


p
5 0.1
5 0.4
10 0.4
f(x)
x
0.12
0.10
0.08
0.06
0.04
0.02
0
0
20
40
60
80
100
120



If X is a negative binomial random variable with parameters p and r,

$$\mu = E(X) = r/p \qquad \text{and} \qquad \sigma^2 = V(X) = r(1-p)/p^2$$

$$X = \sum_{k=1}^{r} X_k, \quad E[X] = \sum_{k=1}^{r} E[X_k] = \sum_{k=1}^{r} \frac{1}{p} = \frac{r}{p}$$

- A binomial random variable is a count of the number of successes in *n* Bernoulli trials.
  - ☞The number of trials is predetermined, and the number of successes is random.
- A negative binomial random variable is a count of the number of trials required to obtain *r* successes.
  - ☞ The number of successes is predetermined, and the number of trials is random.

3.8 Hypergeometric Distribution

- Ex. 3-8

Let X equal the number of nonconforming parts in the sample. Then

$$P(X = 0) = P(\text{both parts conform}) = (800/850)(799/849) = 0.886$$

$$P(X = 1) = P(\text{first part selected conforms and the second part selected does not, or the first part selected does not and the second part selected conforms})$$

$$= (800/850)(50/849) + (50/850)(800/849) = 0.111$$

$$P(X = 2) = P(\text{both parts do not conform}) = (50/850)(49/849) = 0.003$$

- Definition

A set of N objects contains

K objects classified as successes

$N - K$ objects classified as failures

A sample of size n objects is selected randomly (without replacement) from the N objects, where $K \leq N$ and $n \leq N$.

Let the random variable X denote the number of successes in the sample. Then X is a **hypergeometric random variable** and

$$f(x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} \qquad x = \max\{0, n + K - N\} \text{ to } \min\{K, n\} \qquad (3\text{-}13)$$

If X is a hypergeometric random variable with parameters N, K, and n, then

$$\mu = E(X) = np \qquad \text{and} \qquad \sigma^2 = V(X) = np(1-p)\left(\frac{N-n}{N-1}\right) \tag{3-14}$$

where $p = K/N$.

- Ex3-8 Four parts are selected.
  - Four parts do not conform $P(X=4)$
  - What is the probability that two or more parts in the sample do not conform? $P(X \geq 2)$
  - What is the probability that at least one part does not conform.

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

- Finite population correction factor

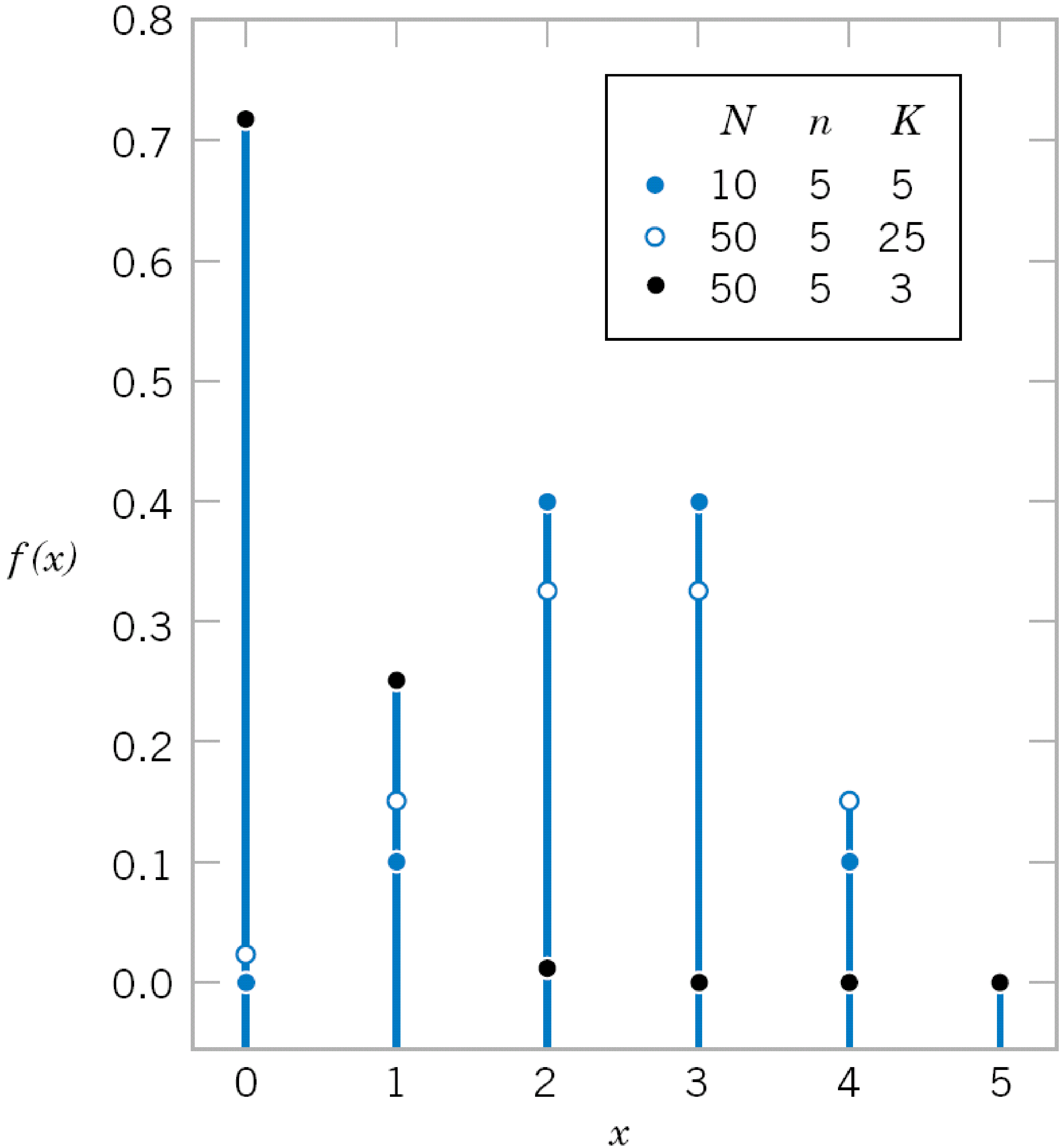
The term in the variance of a hypergeometric random variable

$$\frac{N-n}{N-1}$$

is called the finite population correction factor.

1. The finite population correction represents the correction to the binomial variance that results because the sampling is without replacement from the finite set of size *N*.

2. If *n* is small relative to *N*, the correction is small and the hypergeometric distribution is similar to the binomial.



0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
f(x)
N n K
10 5 5
50 5 25
50 5 3
0 1 2 3 4 5
x

3.9 Poisson Distribution

- Ex3-30
  - n bits, X:the number of bits in error.
  - When the probability that a bit is in error is constant and the transmissions are independent, X has a binomial distribution.
  - Let p denote the probability that a bit is in error.
  - Let $\lambda = pn$. Then, $E[x] = pn = \lambda$ and

$$P(X = x) = \binom{n}{x} P^x (1 - p)^{n-x} = \binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x}$$

- Now, suppose that the number of bits transmitted increases and the probability of an error decreases exactly enough that *pn* remains equal to a constant.
- That is, *n* increases and *p* decreases accordingly, such that $E(X) = \lambda$ remains constant.
- $\lim_{n \to \infty} P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \qquad x = 0, 1, 2, \ldots$
- The number of bits transmitted tends to infinity.
- The range of *X* is the integers from zero to infinity.

- Definition
  - In general, consider an interval T of real numbers partitioned into subintervals of small length Δt and assume that as Δt tends to zero,

(1) the probability of more than one count in a subinterval is zero,

(2) the probability of one count in a subinterval tends to $\lambda \Delta t/T$,

(3) the event in each subinterval is independent of other subintervals.

A random experiment with these properties is called a **Poisson process.**

  - These assumptions imply that the subintervals can be thought of as approximate independent Bernoulli trials with success probability $p = \lambda \Delta t/T$ and the number of trials equal to $n = T/\Delta t$. Here, $pn = \lambda$, and as Δt tends to zero, n tends to infinity.

The random variable X that equals the number of counts in the interval is a **Poisson random variable** with parameter $0<\lambda$, and the probability mass function of X is

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!} \qquad x = 0, 1, 2, \ldots$$

- If X is a Poisson random variable with parameter λ, then

$$\mu = E(X) = \lambda \qquad \text{and} \qquad \sigma^2 = V(X) = \lambda$$

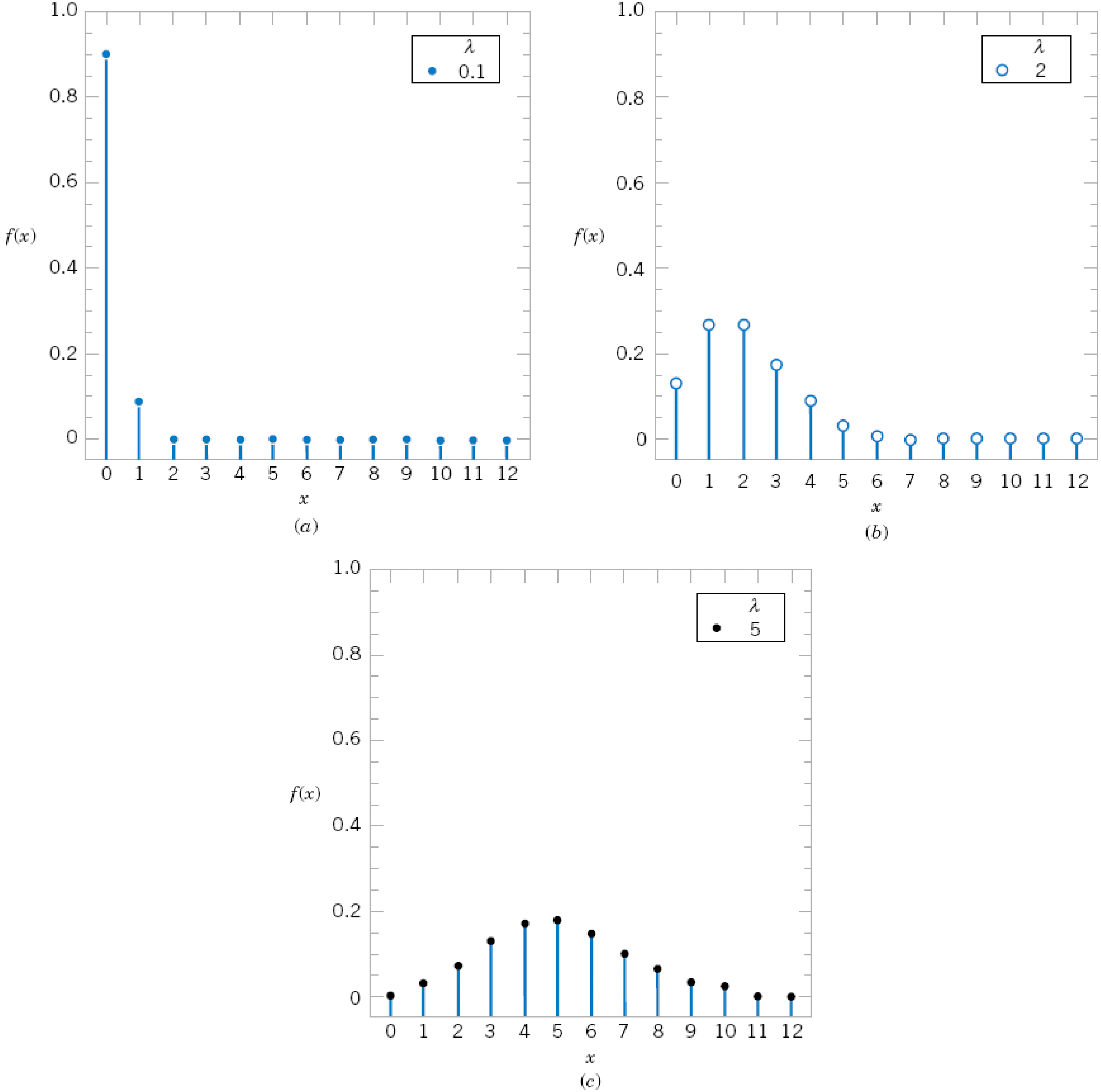
- Ex.3-158

(a) $f(5)=$ exp(-5)(5^5)/5!

(b) mean= 5/h=7.5/(1.5h), $f(10)$=exp(-7.5)(7.5^10)/10!

(c) mean= 2.5/(0.5h),

$P(X<2)=$ $f(0)+f(1)$=exp(-2.5)+exp(-2.5)(2.5)


λ
0.1
2
5
f(x)
x
1.0
0.8
0.6
0.4
0.2
0
0 1 2 3 4 5 6 7 8 9 10 11 12
(a)
(b)
(c)