# MAT 271E: PROBABILITY AND STATISTICS

PROF. DR. CANAN SARICAM

# WEEK 8

## COMMON PROBABILITY DISTRIBUTION FUNCTIONS

# COMMON PROBABILITY DISTRIBUTIONS FOR DISCRETE RANDOM VARIABLE

- Discrete Uniform Distribution
- Bernoulli process
- Binomial probability distribution
- Poisson probability distribution

# DISCRETE UNIFORM DISTRIBUTION

**Definition**

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

$$p(x_i) = 1/n$$

# DISCRETE UNIFORM DISTRIBUTION

EXAMPLE 5.17

The first digit of a part's serial number is equally likely to be any one of the digits 0 through 9. If one part is selected from a large batch and $X$ is the first digit of the serial number, then $X$ has a discrete uniform distribution with probability 0.1 for each value in $R = \{0, 1, 2, \ldots, 9\}$. That is,

$$p(x) = 0.1$$

for each value in $R$. The probability mass function of $X$ is shown in Fig. 5.2

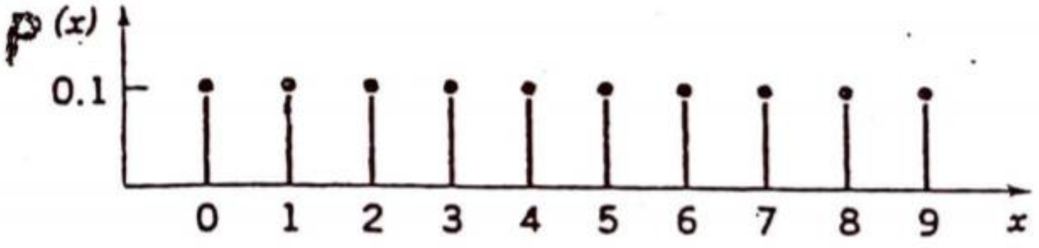


Figure 5.2 Probability mass function for the discrete uniform random variable in Example 5.17

# DISCRETE UNIFORM DISTRIBUTION

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

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

The variance of $X$ is

$$\sigma^2 = \frac{(b-a+1)^2 - 1}{12} \tag{3-6}$$

# DISCRETE UNIFORM DISTRIBUTION

## Example 5.18

EXAMPLE 5.18

We assume that the number of voice lines that are in use at a particular time is a discrete uniform random variable, denoted as $X$. Then,

# DISCRETE UNIFORM DISTRIBUTION

## Example 5.18 - solution

EXAMPLE 5.18

We assume that the number of voice lines that are in use at a particular time is a discrete uniform random variable, denoted as $X$. Then,

$$E(X) = (0 + 48)/2 = 24$$

and

$$\sigma_X = \{[(48 - 0 + 1)^2 - 1]/12\}^{1/2} = 14.14$$

# BINOMIAL PROBABILITY DISTRIBUTION BERNOULLI PROCESS

- A bernoulli random trial is a random trial that has two basic outcomes of qualitative nature.
- To quantify these outcomes, we arbitrarily assign one outcome the value 1 and the other the value zero.
- We shall let p denote the probability that the outcome takes value 1. It then follows that the probability of the outcome 0 is 1-p.
- We shall let B denote the random variable associated with a Bernoulli random trial.
- We have then that $P(B=1)=p$ and $P(B=0)=1-p$
- The probability distribution of a Bernoulli random variable B is called a Bernoulli probability distribution and has the form

$$P(B=b)=p^{b}(1-p)^{1-b};\ b=0.1$$

# BINOMIAL PROBABILITY DISTRIBUTION
# BERNOULLI PROCESS

## Example 5.20

The calculation of a payroll check may be correct or incorrect. Let $B = 1$ correspond to an incorrectly calculated check and let $B = 0$ to an correctly calculated one. If the probabiility of an incorrectly calculated check is $p = 0.05$, then the Bernoulli probability distribution for the random variable B is:

# BINOMIAL PROBABILITY DISTRIBUTION
# BERNOULLI PROCESS

## Example 5.20

EXAMPLE(5.20) :

The calculation of a payroll check may be correct or incorrect. Let B = 1 correspond to an incorrectly calculated check and let B= 0 to an correctly calculated one. If the probabiility of an incorrectly calculated check is p= 0.05, then the Bernoulli probability distribution for the random variable B is:

| b | p(b) |
|---|---|
| 0 | 0.95 |
| 1 | 0.05 |
| Total | 1.00 |

# BINOMIAL PROBABILITY DISTRIBUTION
# BERNOULLI PROCESS

## Bernoulli Process

In statistical applications, we are seldom interested in a single Bernoulli random variable. More commonly, a sequence $B_1, B_2, \ldots\ldots, B_n$ of Bernoulli random variables is under consideration.

A **Bernoulli process** is a sequence of **independent and identically distributed** Bernoulli random variables.

# BINOMIAL PROBABILITY DISTRIBUTION

1. Flip a coin 10 times. Let $X$ = number of heads obtained.
2. A worn machine tool produces 1% defective parts. Let $X$ = number of defective parts in the next 25 parts produced.
3. Each sample of air has a 10% chance of containing a particular rare molecule. Let $X$ = the number of air samples that contain the rare molecule in the next 18 samples analyzed.
4. 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.
5. A multiple choice test contains 10 questions, each with four choices, and you guess at each question. Let $X$ = the number of questions answered correctly.
6. In the next 20 births at a hospital, let $X$ = the number of female births.
7. Of all patients suffering a particular illness, 35% experience improvement from a particular medication. In the next 100 patients administered the medication, let $X$ = the number of patients who experience improvement.

# BINOMIAL PROBABILITY DISTRIBUTION

Repeated Bernoulli process has the probability distribution which is known as binomial probability distribution.

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

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

# BINOMIAL PROBABILITY DISTRIBUTION

## Example 5.22

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 $P(X = 2)$.

# BINOMIAL PROBABILITY DISTRIBUTION

**Example 5.22**

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 $P(X = 2)$.

**$P(X=2) = {}_4C_2\,(0.1)^2(0.9)^2$**

**$P(X=2) = 0{,}0486$**

# BINOMIAL PROBABILITY DISTRIBUTION

**Example 5.23**

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 proba-

a.Find the probability that in next 18 samples, exactly 2 contain pollutant

b.Determine the probability at least 4 samples contain the pollutant

c.Determine the probability that $3 \leq x < 7$

# BINOMIAL PROBABILITY DISTRIBUTION

**Example 5.23**

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 proba-

a.Find the probability that in next 18 samples, exactly 2 contain pollutant

Let $X$ = the number of air samples that contain the rare molecule in the next 18 samples analyzed. Then $X$ is a binomial random variable with $p = 0.1$ and $n = 18$. Therefore,

$$P(X = 2) = \binom{18}{2}(0.1)^2(0.9)^{16}$$

Now $\binom{18}{2} = (18!/[2!\ 16!]) = (18(17)/2) = 153$. Therefore,

$$P(X = 2) = 153(0.1)^2(0.9)^{16} = 0.284$$

# BINOMIAL PROBABILITY DISTRIBUTION

## Example 5.23

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 proba-

b.Determine the probability at least 4 samples contain the pollutant

$$P(X \geq 4) = \sum_{x=4}^{18} \binom{18}{x} (0.1)^x (0.9)^{18-x}$$

However, it is easier to use the complementary event,

$$P(X \geq 4) = 1 - P(X < 4) = 1 - \sum_{x=0}^{3} \binom{18}{x} (0.1)^x (0.9)^{18-x}$$

$$= 1 - [0.150 + 0.300 + 0.284 + 0.168] = 0.098$$

# BINOMIAL PROBABILITY DISTRIBUTION

**Example 5.23**

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 proba-

c.Determine the probability that $3 \leq x < 7$

Determine the probability that $3 \leq X < 7$. Now

$$P(3 \leq X < 7) = \sum_{x=3}^{6} \binom{18}{x} (0.1)^x (0.9)^{18-x}$$
$$= 0.168 + 0.070 + 0.022 + 0.005$$
$$= 0.265$$

# BINOMIAL PROBABILITY DISTRIBUTION

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

$$\mu_X = E(X) = np \quad \text{and} \quad \sigma_X^2 = V(X) = np(1-p)$$

**Example 5.24**

EXAMPLE 5-24

If a worn machine tool produces 1% defective parts and the parts produced are independent, then the mean number of defective parts out of 25 is

# BINOMIAL PROBABILITY DISTRIBUTION

## Example 5.24

EXAMPLE 5-24 ........................................................................

If a worn machine tool produces 1% defective parts and the parts produced are independent, then the mean number of defective parts out of 25 is

$$E(X) = 25(0.01) = 0.25$$

The variance of the number of defective parts is

$$V(X) = 25(0.01)(0.99) = 0.2475$$

# BINOMIAL PROBABILITY DISTRIBUTION

**Example 5.25**

A quality control engineer wants to check whether (in accordance with specifications) the components shipped by her company are in good working condition. Prior to shipment, she randomly selects 15 components from very large lot. She allows the lot to be shipped only if all 15 components are in good working condition. Otherwise, each of the components in the lot is tested, and any bad components are replaced with good components.

a.What is the probability that the engineer will commit the error of holding in a lot for further inspection even though 95% of the components are in good working condition.

b.What is the probability that the engineer will commit the error of shipping a lot without further inspection when only 90% of the components are in good working condition

# BINOMIAL PROBABILITY DISTRIBUTION

**Example 5.25**

*Solution*

(a) We want the probability that there are one or more bad components in the sample of 15 components when the probability of a good component is 0.95. This is exactly the same as 1 minus the probability that all 15 of the components are good; i.e., 1 minus the probability of 15 successes in 15 trials. If we define a success as a good component, $p = 0.95$. Therefore, p(one or more bad components) is equal to

$$1 - P(\text{no failures}) = 1 - p(15) = 1 - {}_{15}C_{15}(0.95)^{15}(0.05)^{0} = 1 - (0.95)^{15} = 0.5367$$

If we define a "success" as a *bad* component, $p = 0.05$; and p(one or more bad components) is equal to

$$1 - P(\text{no successes}) = 1 - {}_{15}C_{0}(0.05)^{0}(0.95)^{15} = 1 - (0.95)^{15} = 0.5367$$

(b) We want the probability that no bad components are present in the sample of 15 components when the probability of a good component is 0.90. That probability is equal to the probability that all of the components are good. $p(15) = {}_{15}C_{15}(0.9)^{15}(0.1)^{0} = (0.9)^{15} = 0.205891.$

# POISSON DISTRIBUTION

Poisson distribution is used for the random phenomena dealing with **the number of events occurring in a fixed time interval.**

Examples:

- The number of machines, X in a plant that break down during a day is poisson random variable. Management wishes to know how likely it is that there are no breakdown in a day.
- The number of units X of an item sold from stock during a week is a Poisson random variable. The inventory controller wishes to know the probability that more than five units will be sold in a week
- The number of persons X arriving at an automatic bank teller between 9 AM and 10 AM is a poisson random variable. A system consultant needs to know the probability that more than 10 persons will arrive during the one hour period.
- The poisson probability distribution also applies to random occurrances that are not related directly to time such as the number of typographical error on a page or the number of faults on a fabric per square meter.

# POISSON DISTRIBUTION

- A poisson random variable is  discrete random variable that can take on any integer value from 0 to infinity.

The *Poisson distribution*

$$p(x) = \frac{\lambda^x e^{-\lambda}}{x!} \qquad x = 0, 1, 2, \ldots \qquad e = 2.71828$$

- The poisson distribution is often used to characterize physical situations in which the number of events during a specific period of time of interest, like the number of customers arriving at a bank during one hour.
- The random variable in the equation is the number of occurences X, and $\lambda > 0$ is the rate parameter- the average number of occurences in a specific time period.
- In poisson distribution, we assume that occurences are equally likely to happen during any time interval and that one occurence has no effect on the probability of another.

# POISSON DISTRIBUTION

Given an interval of real numbers, assume counts occur at random throughout the interval. If the interval can be partitioned into subintervals of small enough length such that

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

(2) the probability of one count in a subinterval is the same for all subintervals and proportional to the length of the subinterval, and

(3) the count in each subinterval is independent of other subintervals, the random experiment is called a **Poisson process.**

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

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!} \quad x = 0, 1, 2, \ldots \tag{3-15}$$

# POISSON DISTRIBUTION

**Example 5.27**

The number of customer orders for 21 inch microcomputer monitors at a large mail order firm averages 20 monitors per week. The average level of customer demand is constant and customers do not affect one another in their buying habits. Find the probability that exactly 17 monitors will be purchased in  week.

# POISSON DISTRIBUTION

**Example 5.27**

The number of customer orders for 21 inch microcomputer monitors at a large mail order firm averages 20 monitors per week. The average level of customer demand is constant and customers do not affect one another in their buying habits. Find the probability that exactly 17 monitors will be purchased in  week.

λ=20

$$P(X=17)= \frac{20^{17}e^{-20}}{17!}$$

P(X=17)=0.07595

# POISSON DISTRIBUTION

**Example 5.30**

For the case of the thin copper wire, suppose that the number of flaws follows a Poisson distribution with a mean of 2.3 flaws per millimeter. Determine the probability of exactly 2 flaws in 1 millimeter of wire. Determine the probability of 10 flaws in 5 millimeters of wire. Determine the probability of at least 1 flaw in 2 millimeters of wire.

# POISSON DISTRIBUTION

## Example 5.30

Let $X$ denote the number of flaws in 1 millimeter of wire. Then, $E(X) = 2.3$ flaws and

$$P(X = 2) = \frac{e^{-2.3}2.3^2}{2!} = 0.265$$

Determine the probability of 10 flaws in 5 millimeters of wire. Let $X$ denote the number of flaws in 5 millimeters of wire. Then, $X$ has a Poisson distribution with

$$E(X) = 5 \text{ mm} \times 2.3 \text{ flaws/mm} = 11.5 \text{ flaws}$$

Therefore,

$$P(X = 10) = e^{-11.5} \frac{11.5^{10}}{10!} = 0.113$$

Determine the probability of at least 1 flaw in 2 millimeters of wire. Let $X$ denote the number of flaws in 2 millimeters of wire. Then, $X$ has a Poisson distribution with

$$E(X) = 2 \text{ mm} \times 2.3 \text{ flaws/mm} = 4.6 \text{ flaws}$$

Therefore,

$$P(X \geq 1) = 1 - P(X = 0) = 1 - e^{-4.6} = 0.9899$$

# POISSON DISTRIBUTION

**Example 5.31**

Contamination is a problem in the manufacture of optical storage disks. The number of particles of contamination that occur on an optical disk has a Poisson distribution, and the average number of particles per centimeter squared of media surface is 0.1. The area of a disk under study is 100 squared centimeters. Find the probability that 12 particles occur in the area of a disk under study. Find the probability that zero particles occur in the area of the disk under study. Determine the probability that 12 or fewer particles occur in the area of the disk under study.

# POISSON DISTRIBUTION

## Example 5.31

Let $X$ denote the number of particles in the area of a disk under study. Because the mean number of particles is 0.1 particles per $cm^2$

$$E(X) = 100 \text{ cm}^2 \times 0.1 \text{ particles/cm}^2 = 10 \text{ particles}$$

Therefore,

$$P(X = 12) = \frac{e^{-10}10^{12}}{12!} = 0.095$$

The probability that zero particles occur in the area of the disk under study is

$$P(X = 0) = e^{-10} = 4.54 \times 10^{-5}$$

Determine the probability that 12 or fewer particles occur in the area of the disk under study. The probability is

$$P(X \le 12) = P(X = 0) + P(X = 1) + \cdots + P(X = 12) = \sum_{i=0}^{12} \frac{e^{-10}10^i}{i!}$$

Because this sum is tedious to compute, many computer programs calculate cumulative Poisson probabilities. From one such program, $P(X \le 12) = 0.791$.

# POISSON APPROXIMATION TO BINOMIAL DISTRIBUTION

Poisson probabilities can be used to approximate binomial probabilities when n is large and p s small.

As a rule of thumb, this approximation is acceptable if $n \geq 20$ and $p \leq 0.05$. If $n \geq 100$ and $np \leq 10$, it is usually an excellent approximation.

To make poisson approximation fo binomial distribution, $\lambda = np$

# POISSON APPROXIMATION TO BINOMIAL DISTRIBUTION

**Example**

It is known that 3% of the circuit boards from a production line are defective. If a random sample of 120 circuit boards is taken from this production line, use the poisson approximation to estimate the probability that the sample contains

a. Exactly 2 defective boards

b. At least 2 defective boards

In this case, $n \geq 100$ and $np \leq 10$  $120 \geq 100$,  $120*0.03=3.6 \leq 10$

So, poisson distribution can aproximate binomial distribution $\lambda=120*0.03=3.6$

# POISSON APPROXIMATION TO BINOMIAL DISTRIBUTION

**Binomial distribution**

$P(X=2) = {}_{120}C_2\,(0.03)^2(0.97)^{118}$

$P(X=2) = 0.175$

**Poisson distribution**

$\lambda=3.6$

$P(X=2)=\frac{e^{-3.6}3.6^2}{2!}$

$P(X=2)=0.177$

# POISSON APPROXIMATION TO BINOMIAL DISTRIBUTION

**Binomial distribution**

$P(X \geq 2) = 1 - [P(X=0) + P(X=1)]$

$P(X \geq 2) = 1 - [{}_{120}C_0\,(0.03)^0(0.97)^{120}$

$+\ {}_{120}C_1\,(0.03)^1(0.97)^{119}]$

$P(X \geq 2) = 0.878$

**Poisson distribution**

$\lambda = 3.6$

$P(X \geq 2) = 1 - [P(X=0) + P(X=1)]$

$P(X \geq 2) = 1 - [\frac{e^{-3.6}3.6^0}{0!} + \frac{e^{-3.6}3.6^1}{1!}]$

$P(X=2) = 0.876$