# Lecture 6.

# Examples of probability mass functions (discrete variables)

prof. dr hab.inż. Katarzyna Zakrzewska
Katedra Elektroniki, AGH
e-mail: zak@agh.edu.pl
http://home.agh.edu.pl/~zak

# Outline:

- Definitions of mean and variance for discrete variables
- Discrete uniform distribution
- Binomial (Bernoulli) distribution
- Geometric distribution
- Poisson distribution

# MEAN AND VARIANCE OF A DISCRETE RANDOM VARIABLE

**Definition**

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

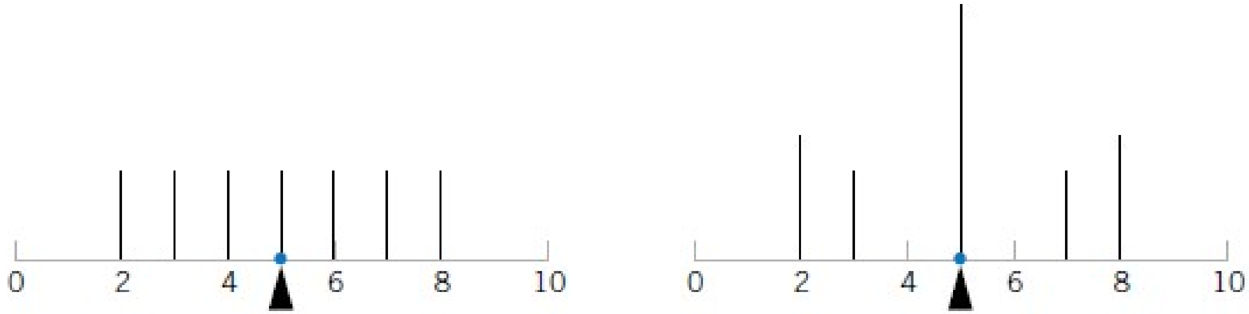
$$\mu = E(X) = \sum_x x f(x) \tag{3-3}$$

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

$$\sigma^2 = V(X) = E(X-\mu)^2 = \sum_x (x-\mu)^2 f(x) = \sum_x x^2 f(x) - \mu^2$$

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

Mean and variance are two measures that do not uniquely identify a probability distribution. Below you can find two different distributions that have the same mean and variance.



# MEAN AND VARIANCE OF A DISCRETE RANDOM VARIABLE

The variance of a random variable X can be considered to be the **expected value** of a specific function of X:

$$h(X) = (X - \mu)^2$$

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 $f(x)$,

$$E[h(X)] = \sum_x x h(x) f(x) \tag{3-4}$$

# Discrete uniform distribution

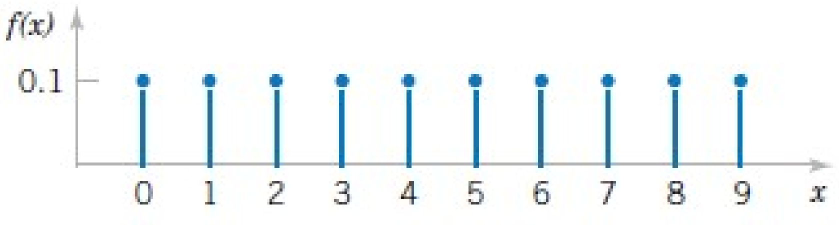
The simplest discrete random variable is one that assumes only a finite number of possible values, each with equal probability.

A random variable X that assumes each of the values $x_1, x_2, \ldots, x_n$ with equal probability 1/n, is frequently of interest.

**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,

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



# Discrete uniform distribution

Suppose the range of the discrete random variable X is the consecutive integers: a, a+1, a+2,….b for a≤b.

The range of X contains b-a+1 values each with probability 1/(b-a+1).

According to definition mean value equals to:

$$\mu = \sum_{k=a}^{b} k\left(\frac{1}{b - a + 1}\right)$$

The algebraic identity $\sum_{k=a}^{b} k = \frac{b(b+1) - (a-1)a}{2}$ can be used to simplify the result to $\mu = (b + a)/2$. The derivation of the variance is left as an exercise.

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

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

# Examples of probability distributions – discrete variables

**Two-point distribution** (zero-one), e.g. coin toss, head = failure x=0, tail = success x=1,  p – probability of success, its distribution:

| $x_i$ | 0 | 1 |
|---|---|---|
| $p_i$ | 1-p | p |

**Binomial** (Bernoulli)

$$p_k = \binom{n}{k} \cdot p^k (1-p)^{n-k} \quad , \; k = 0, 1, \ldots, n$$

where 0<p<1; X={0, 1, 2, … k} k – number of successes when n-times sampled with replacement

For k=1  two-point distribution

# Examples of probability distributions – discrete variables

Consider the following random experiments and random variables:

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 distribution

**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

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

# Binomial distribution - assumptions

**Random experiment consists of *n* Bernoulli trials :**

1. **Each trial <u>is independent of others.</u>**
2. **Each trial can have only two results: „success" and „failure" (binary!).**
3. **Probability of success *p* is <u>constant.</u>**

**Probability $p_k$ of an event that random variable X will be equal to the number of k-successes at n trials.**

$$p_k = \binom{n}{k} \cdot p^k (1-p)^{n-k} \quad , \; k = 0, 1, \ldots, n$$

# Pascal's triangle

Symbol $\binom{n}{k} = \frac{n!}{(n-k)!k!}$

$n = 0$ $\binom{0}{0} = 1$

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

$n = 1$ $\binom{1}{0} = 1$ $\binom{1}{1} = 1$

Newton's binomial

$n = 2$ $\binom{2}{0} = 1$ $\binom{2}{1} = 2$ $\binom{2}{2} = 1$

# Pascal's triangle

| | | |
|---|---|---|
| 1 | | n = 0 |
| 1 + 1 | | n = 1 |
| 1 2 1 | | n = 2 |
| 1 3 3 1 | | n = 3 |
| 1 4 6 4 1 | | n = 4 |
| 1 5 10 10 5 1 | | n = 5 |
| 1 6 15 20 15 6 1 | | n = 6 |

# Bernoulli distribution

**Example 6.1**

Probability that in a company the daily use of water will not exceed a certain level is p=3/4. We monitor a use of water for 6 days.

Calculate a probability the daily use of water will not exceed the set-up limit in 0, 1, 2, …, 6 consecutive days, respectively.

**Data:**

$$p = \frac{3}{4} \qquad q = \frac{1}{4} \qquad N = 6 \qquad k = 0, 1, \ldots, 6$$

# Bernoulli distribution

$$k = 0 \qquad P(k=0) = \binom{6}{0} \cdot \left(\frac{3}{4}\right)^{0} \cdot \left(\frac{1}{4}\right)^{6}$$

$$k = 1 \qquad P(k=1) = \binom{6}{1} \cdot \left(\frac{3}{4}\right)^{1} \cdot \left(\frac{1}{4}\right)^{5}$$

$$k = 2 \qquad P(k=2) = \binom{6}{2} \cdot \left(\frac{3}{4}\right)^{2} \cdot \left(\frac{1}{4}\right)^{4}$$

$$k = 3 \qquad P(k=3) = \binom{6}{3} \cdot \left(\frac{3}{4}\right)^{3} \cdot \left(\frac{1}{4}\right)^{3}$$

$$k = 4 \qquad P(k=4) = \binom{6}{4} \cdot \left(\frac{3}{4}\right)^{4} \cdot \left(\frac{1}{4}\right)^{2}$$

$$k = 5 \qquad P(k=5) = \binom{6}{5} \cdot \left(\frac{3}{4}\right)^{5} \cdot \left(\frac{1}{4}\right)^{1}$$

$$k = 6 \qquad P(k=6) = \binom{6}{6} \cdot \left(\frac{3}{4}\right)^{6} \cdot \left(\frac{1}{4}\right)^{0}$$

# Bernoulli distribution

$k=0$ $$P(0)=1\cdot 1\cdot \frac{1}{4^6}\cong 0.00024$$

$k=1$ $$P(1)=6\cdot \frac{3}{4}\cdot \frac{1}{4^5}=\frac{6\cdot 3}{4^6}=18\cdot P(0)\cong 0.004$$

$k=2$ $$P(2)=15\cdot \left(\frac{3}{4}\right)^2\cdot \frac{1}{4^4}=\frac{15\cdot 9}{4^6}=135\cdot P(0)\cong 0.033$$

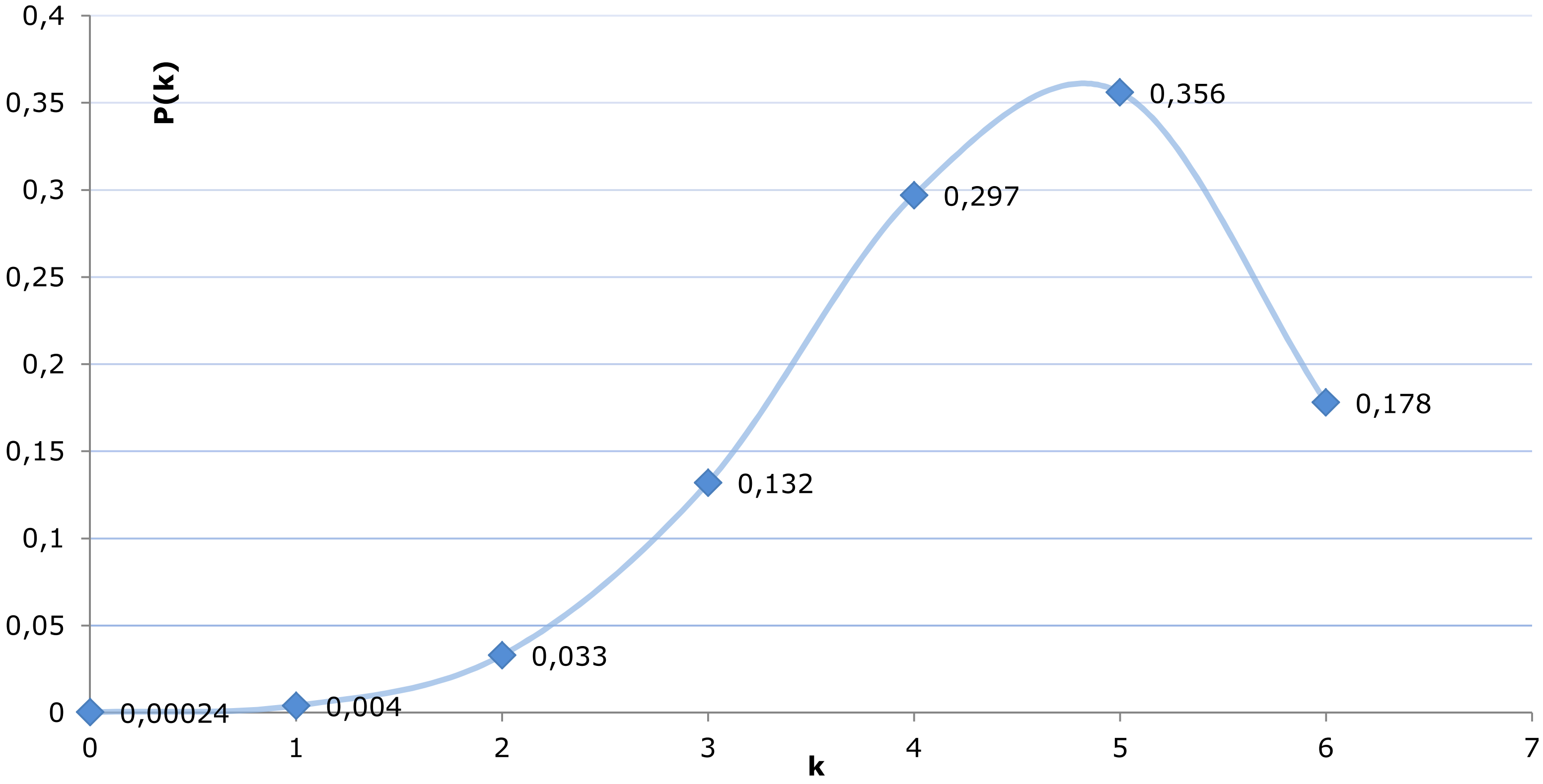
$k=3$ $$P(3)=20\cdot \left(\frac{3}{4}\right)^3\cdot \frac{1}{4^3}=\frac{20\cdot 9\cdot 3}{4^6}=540\cdot P(0)\cong 0.132$$

$k=4$ $$P(4)=15\cdot \left(\frac{3}{4}\right)^4\cdot \frac{1}{4^2}=\frac{15\cdot 9\cdot 9}{4^6}=1215\cdot P(0)\cong 0.297$$

$k=5$ $$P(5)=6\cdot \left(\frac{3}{4}\right)^5\cdot \frac{1}{4^1}=\frac{6\cdot 9\cdot 9\cdot 3}{4^6}=1458\cdot P(0)\cong 0.356$$

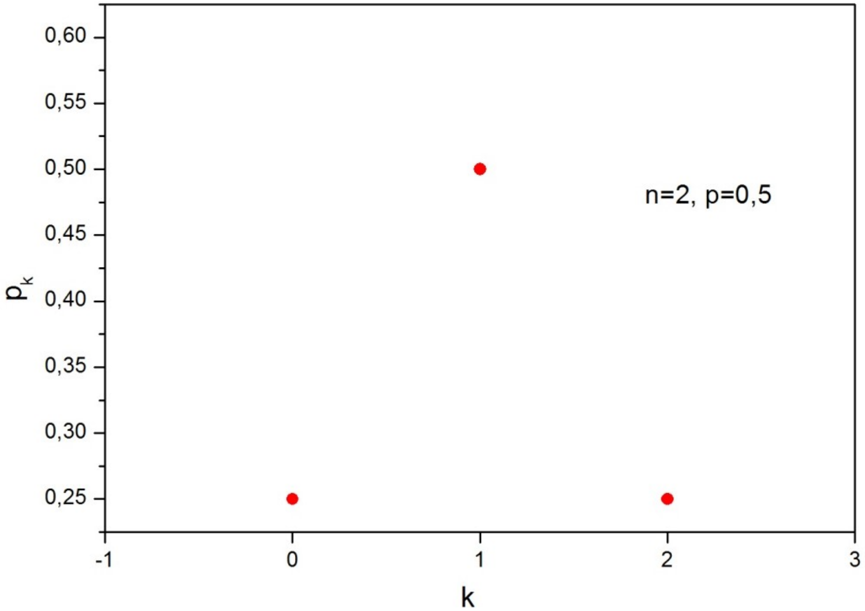
$k=6$ $$P(6)=1\cdot \left(\frac{3}{4}\right)^6\cdot \frac{1}{4^0}=\frac{9\cdot 9\cdot 9}{4^6}=729\cdot P(0)\cong 0.178$$

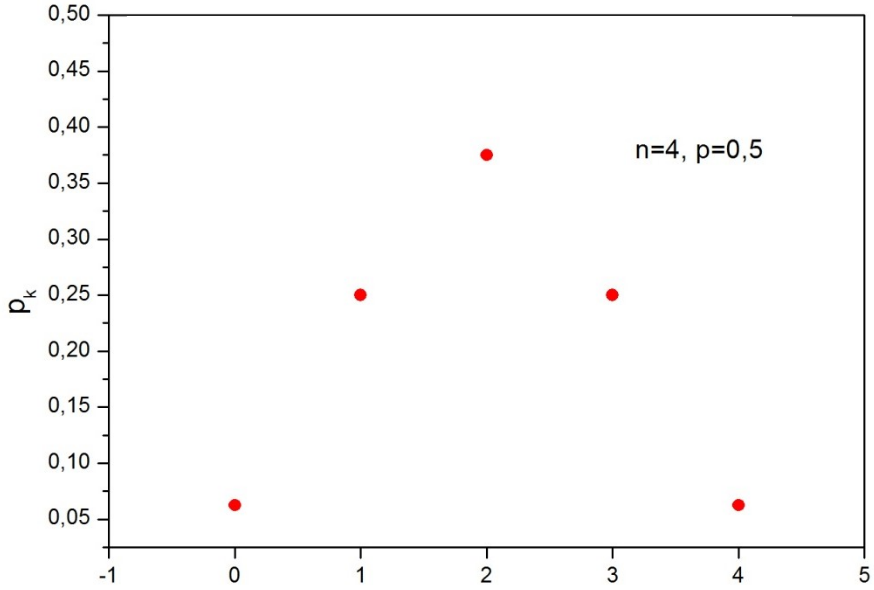
# Bernoulli distribution

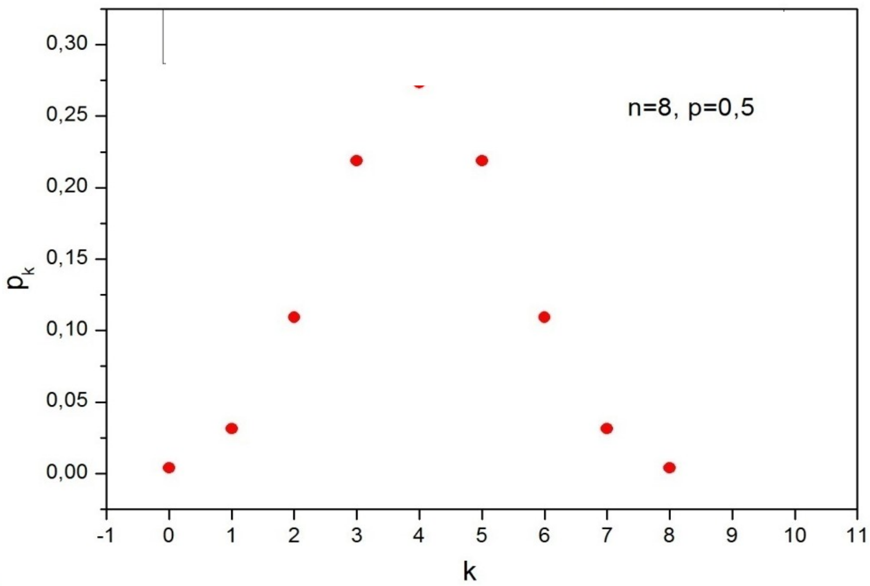


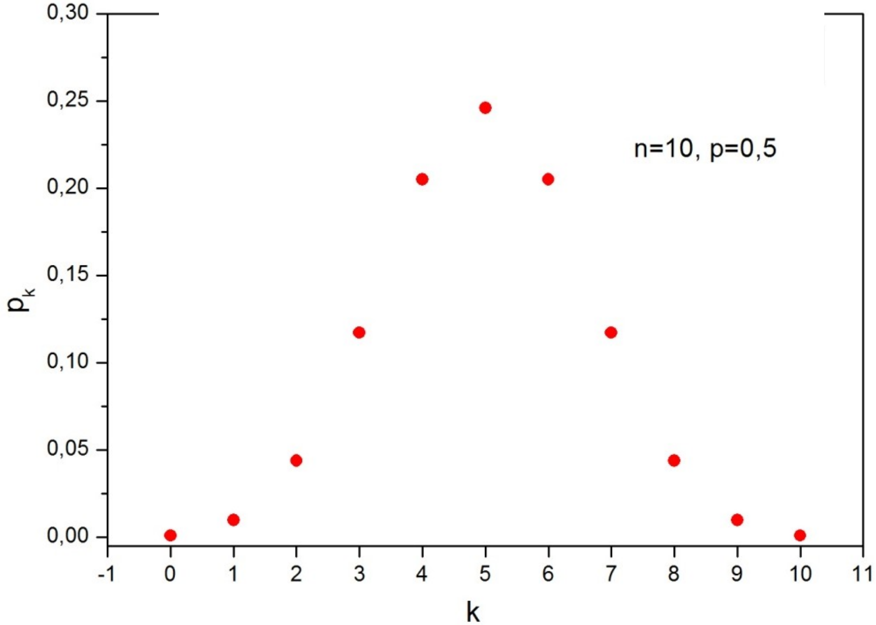

Maximum for k=5

# Bernoulli distribution

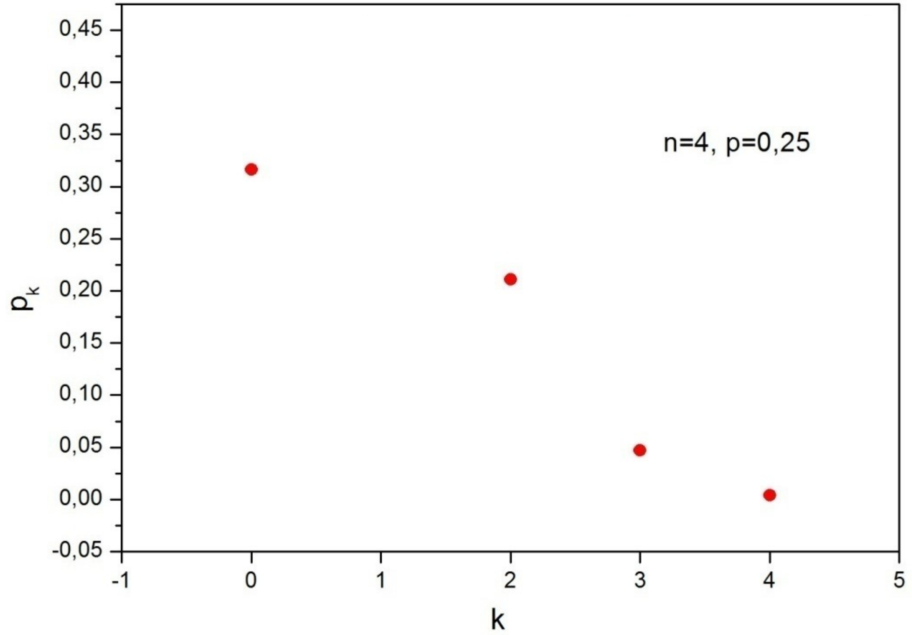


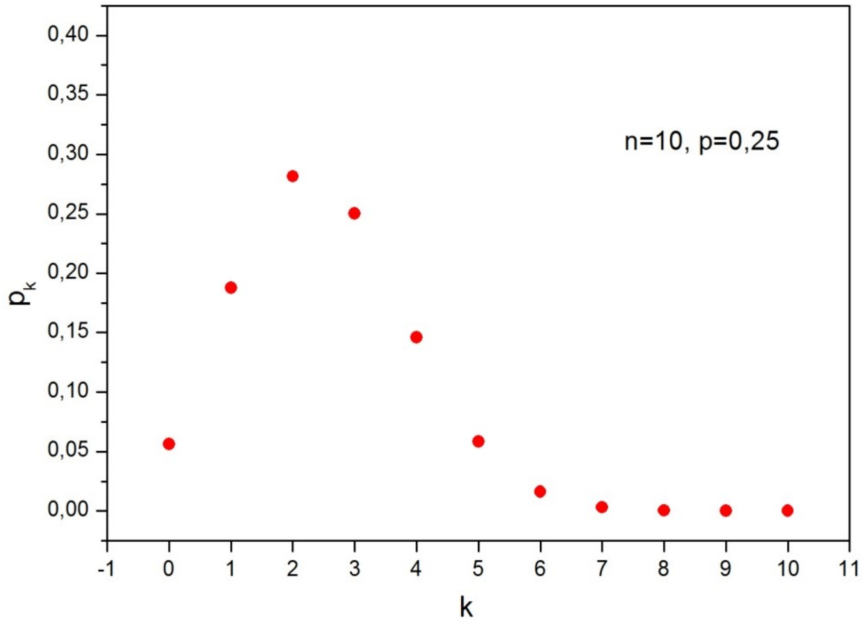



# Bernoulli distribution




Expected value

$$E(X) = \mu = np$$

Variance

$$V(X) = \sigma^2 = np(1-p)$$

# Errors in transmission

**Example 6.2**

Digital channel of information transfer is prone to errors in single bits. Assume that the probability of single bit error is $p=0.1$

Consecutive errors in transmissions are independent. Let X denote the random variable, of values equal to the number of bits in error, in a sequence of 4 bits.

E - bit error, O - no error

OEOE corresponds to $X=2$; for EEOO - $X=2$ (order does not matter)

# Errors in transmission

**Example 6.2 cd**

For X=2 we get the following results:

{EEOO, EOEO, EOOE, OEEO, OEOE, OOEE}

What is a probability of P(X=2), i.e., two bits will be sent with error?

Events are independent, thus

$P(EEOO)=P(E)P(E)P(O)P(O)=(0.1)^2\ (0.9)^2\ = 0.0081$

Events are mutually exhaustive and have the same probability, hence

$P(X=2)=6\ P(EEOO)= 6\ (0.1)^2\ (0.9)^2\ = 6\ (0.0081)=0.0486$

# Errors in transmission

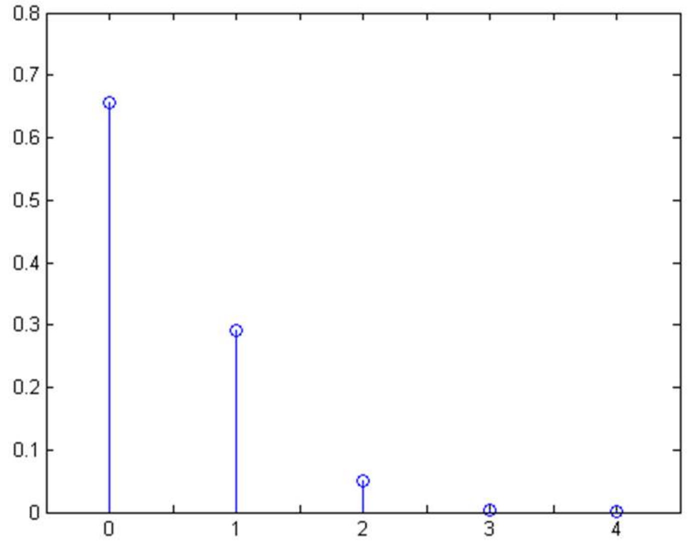
**Example 6.2 continued**

$$\binom{4}{2} = \frac{4\,!}{(2)!2!} = 6$$

Therefore, P(X=2)=6 $(0.1)^2$ $(0.9)^2$ is given by Bernoulli distribution

$$P(X = x) = \binom{4}{x} \cdot p^{x} (1 - p)^{4-x} \;,\; x = 0, 1, 2, 3, 4,\; p = 0.1$$

| |
|---|
| 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 |



# Errors in transmission – calculation of mean and variance

Mean:

$$\mu = E(X) = 0f(0) + 1f(1) + 2f(2) + 3f(3) + 4f(4)$$
$$= 0(0.6561) + 1(0.2916) + 2(0.0486) + 3(0.0036) + 4(0.0001)$$
$$= 0.4$$

Although $X$ never assumes the value 0.4, the weighted average of the possible values is 0.4.
To calculate $V(X)$, a table is convenient.

Variance:

| $x$ | $x - 0.4$ | $(x - 0.4)^2$ | $f(x)$ | $f(x)(x - 0.4)^2$ |
|---|---|---|---|---|
| 0 | −0.4 | 0.16 | 0.6561 | 0.104976 |
| 1 | 0.6 | 0.36 | 0.2916 | 0.104976 |
| 2 | 1.6 | 2.56 | 0.0486 | 0.124416 |
| 3 | 2.6 | 6.76 | 0.0036 | 0.024336 |
| 4 | 3.6 | 12.96 | 0.0001 | 0.001296 |

| |
|---|
| 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 |

$$V(X) = \sigma^2 = \sum_{i=1}^{5} f(x_i)(x_i - 0.4)^2 = 0.36$$

The alternative formula for variance could also be used to obtain the same result.

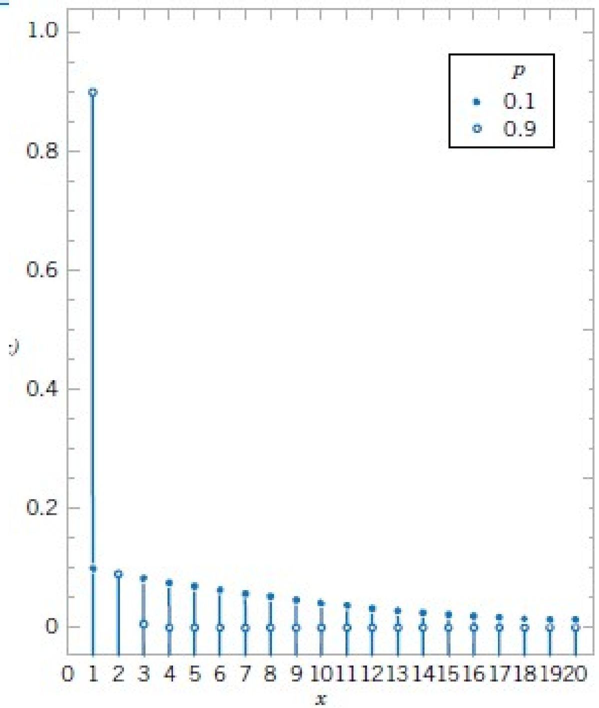
# Geometric distribution

**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 the first success. Then $X$ is a **geometric random variable** with parameter $0 < p < 1$ and

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

The height of the line at x is (1-p) times the height at the line at x-1. That is, the probabilities decrease in a geometric progression. The distribution acquires its name from this result.



# Geometric distribution

If $X$ is a geometric random variable with parameter $p$,

$$\mu = E(X) = 1/p \qquad \text{and} \qquad \sigma^2 = V(X) = (1-p)/p^2 \qquad (3\text{-}10)$$

Lack of memory property (the system will not wear out): A geometric random variable has been defined as the number of trials until the first success. However, because the trials are independent, the count of the number of trials until the next success can be started at any trial without changing the probability distribution of the random variable.

**Example 6.3** In the transmission of bits, if 100 bits are transmitted, the probability that the first error, after bit 100, occurs on bit 106 is the probabability that the next six outcomes are OOOOOE and can be calculated as

$$P(X=6) = p^1(1-p)^5 \ , \ \ p = 0.1$$

This result is identical to the probability that the initial error occurs on bit 6.

# Poisson's distribution

Consider the transmission of n bits over a digital communication channel. Let the random variable X equal the number of bits in error. When the probability that a bit is in error p is constant and the transmissions are independent, X has binomial distribution.

We introduce a parameter λ=pn (E(X) = λ)

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

Let us assume that n increases while p decreases, but λ=pn remains constant. Bernoulli distribution changes to Poisson's distribution.

$$\lim_{n\to\infty} P(X = x) = \lim_{n\to\infty} \binom{n}{x}\left(\frac{\lambda}{n}\right)^{x}\left(1-\frac{\lambda}{n}\right)^{n-x} = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

# Poisson's distribution

It is one of the rare cases where expected value equals to variance:

$$E(X) = np = \lambda$$

Why?

$$V(X) = \sigma^2 = \lim_{n\to\infty, p\to 0} (np - np^2) = np = \lambda$$

# Poisson's distribution

**Definition**

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

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

# Poisson's distribution

Flaws occur at random along the length of a thin copper wire. Let $X$ denote the random variable that counts the number of flaws in a length of $L$ millimeters of wire and suppose that the average number of flaws in $L$ millimeters is $\lambda$.

The probability distribution of $X$ can be found by reasoning in a manner similar to the previous example. Partition the length of wire into $n$ subintervals of small length, say, 1 micrometer each. If the subinterval chosen is small enough, the probability that more than one flaw occurs in the subinterval is negligible. Furthermore, we can interpret the assumption that flaws occur at random to imply that every subinterval has the same probability of containing a flaw, say, $p$. Finally, if we assume that the probability that a subinterval contains a flaw is independent of other subintervals, we can model the distribution of $X$ as approximately a binomial random variable. Because

$$E(X) = \lambda = np$$

we obtain

$$p = \lambda/n$$

That is, the probability that a subinterval contains a flaw is $\lambda/n$. With small enough subintervals, $n$ is very large and $p$ is very small. Therefore, the distribution of $X$ is obtained as in the previous example.

# Poisson's distribution

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.

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,

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

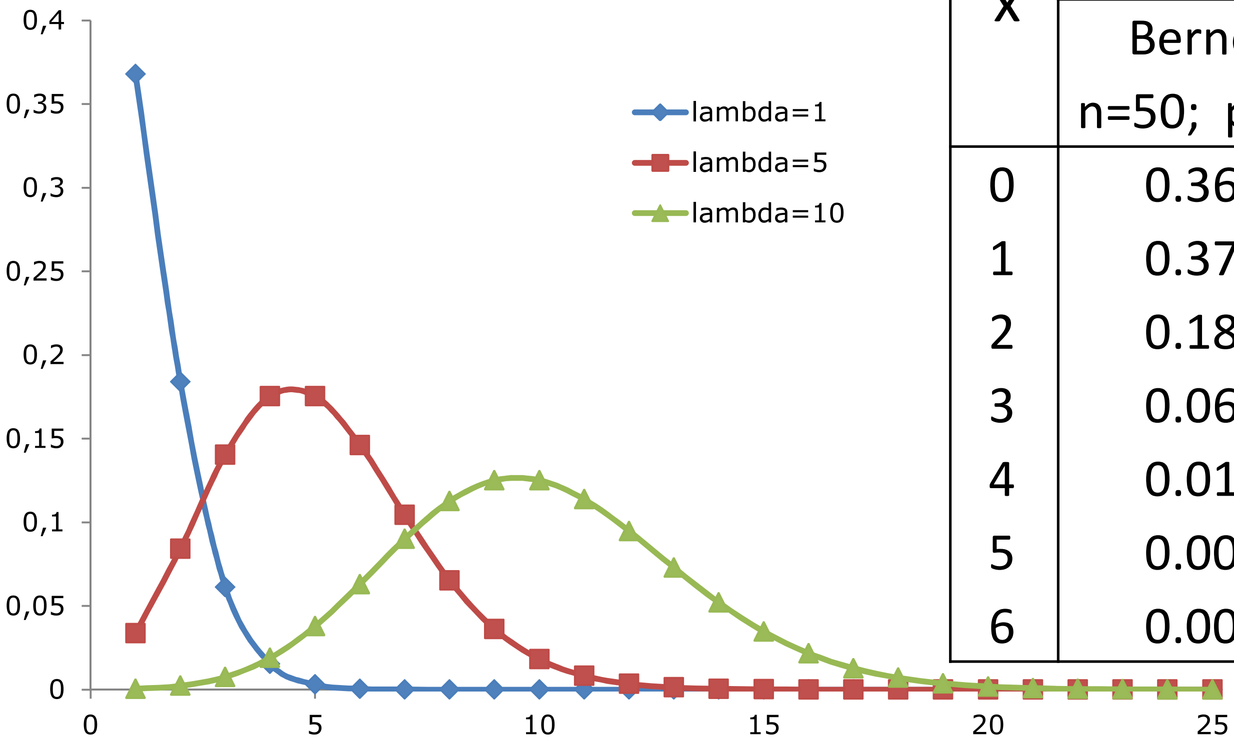
Determine the probability of at least 1 flaw in 2 millimeters of wire. Let $X$ denote the number of flaws in 2 millimeters of wire. Then, $X$ has a Poisson distribution with

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

Therefore,

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

# Poisson's distribution




| x | p(X) | |
|---|---|---|
| | Bernoulli n=50; p=0.02 | Poisson: $\lambda$=1 |
| 0 | 0.364 | 0.368 |
| 1 | 0.372 | 0.368 |
| 2 | 0.186 | 0.184 |
| 3 | 0.061 | 0.061 |
| 4 | 0.014 | 0.015 |
| 5 | 0.003 | 0.003 |
| 6 | 0.000 | 0.001 |

(x- integer, infinite; $x \geq 0$) For big n Bernoulli distribution resembles Poisson's distribution