Slides
Probability Theory
Chapter 4

Ch. 4: Discrete distributions

How to compute probabilities or expectations of (integer) **numbers** of occurring events.

Main concepts in chapters 1-3

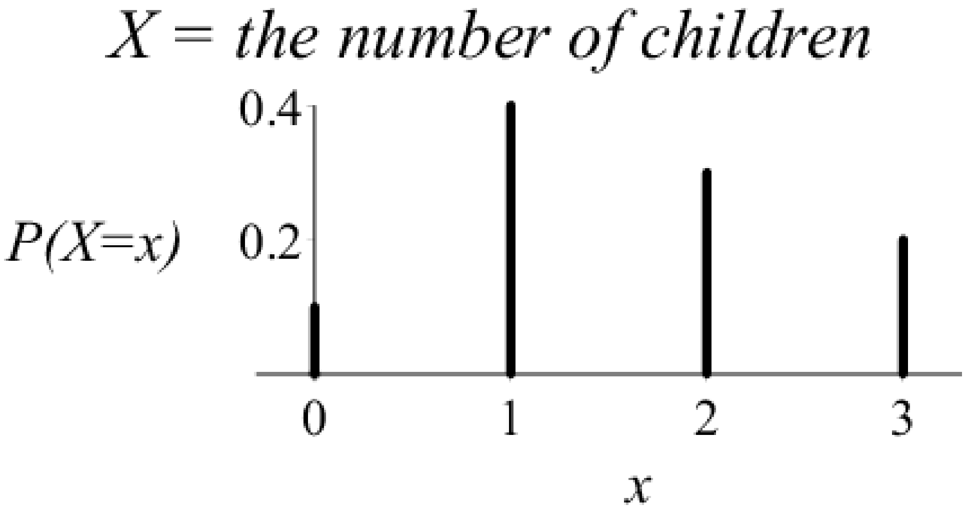
- **Stochastic experiment**
- **Sample space** S
- **Probability measure *P*** assigns a probability $P(A)$ to A.
- **Probability space** $(\boldsymbol{S}, \boldsymbol{P})$
- **Conditional probability:** P is defined on a part B of S: $\big(S, P(.\,|B)\big)$ is a probability space as well.

Ch. 4: **discrete random variables** and their distributions; **expectation and variance.**

Example: X = "the number of children of a woman in a population"

x	0	1	2	3	*total*
$P(X = x)$	0.1	0.4	0.3	0.2	*1*

The **graph of** the distribution: a **bar graph** of probabilities.



Verify that:

1. $P(X = x) \geq 0$

2. $\sum P(X = x) = 1$

(1. and 2. should hold for any discrete distribution)

How can you calculate the average number of children?

Weigh the number of children (x) with the probability $P(X = x)$:

$\sum_x x \cdot P(X = x)$ = 0×0.1 +1×0.4 + 2×0.3 + 3×0.2= 1.6 child/woman

Definition: **a Random Variable (r.v.) *X* :**

a (real) function $X\colon S \to \mathbb{R}$, for an experiment with sample space S, that assigns to **each outcome *s* a real number $\boldsymbol{X(s)}$**.

"A quantitative aspect of the experiment"

Range $\boldsymbol{S_X} = \{X(\boldsymbol{s}) \mid \boldsymbol{s} \in \boldsymbol{S}\}$

The r.v. X is **discrete** if S_X is **denumerable,** so $S_X = \{x_1, \dots, xn\}$ or $S_X = \{x_1, x_2, \dots .\}$

(continuous r.v.: the range is an interval in $\mathbb{R}$)

The **event** $\boldsymbol{X = x}$ is: $\{s \mid X(s) = x\}$

The probability function of X:
$P(X = x)$, where $x \in S_X$

Requirements: **1. $P(X = x) \geq 0$**

2. $\sum_{x \in S_X} P(X = x) = 1$

Distribution of X: a **list** or a **formula** defining the probabilities $P(X = x)$ **for all x** in S_X

Expectation or **Expected value** of X:

$$E(X) = \sum_{x \in S_X} x \cdot P(X = x)$$

$E(X)$ is a **measure of centre** of the distribution: "the weighted average"

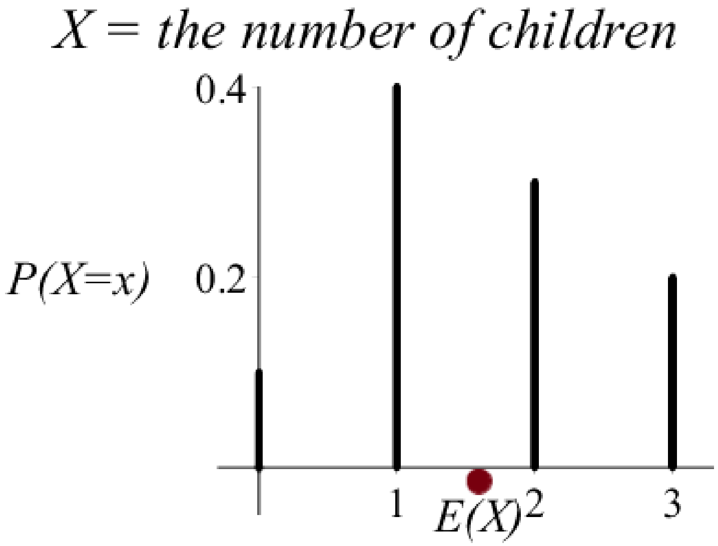
Notation: $E(X) = EX = \mu = \mu_X$

Expectation or **Expected value** of a r.v. X:

$$\boldsymbol{E(X) = \sum_x \; x \cdot P(X = x)}$$

$\boldsymbol{E(X)}$ is a measure of centre:

- **Weighted average**
- **“Population mean”**
- **“Point of balance”** →

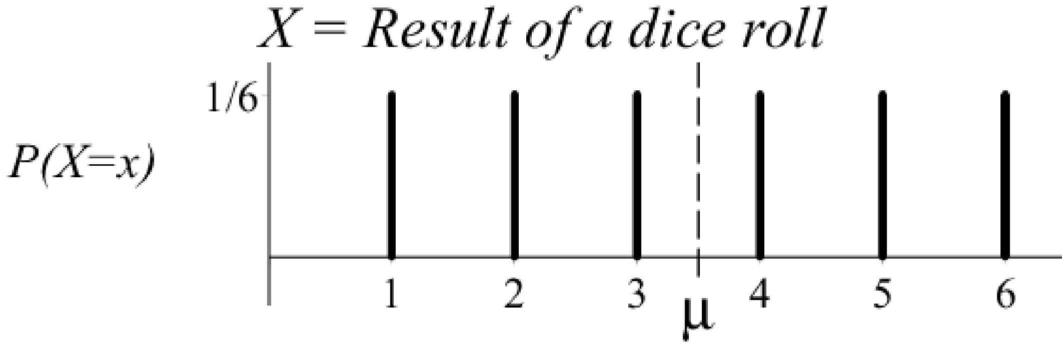


If $\boldsymbol{x = \mu}$ is **the axis of symmetry** of $P(X = x)$, then $\boldsymbol{E(X) = \mu}$**.**

e.g.: X = the result of one roll of a dice

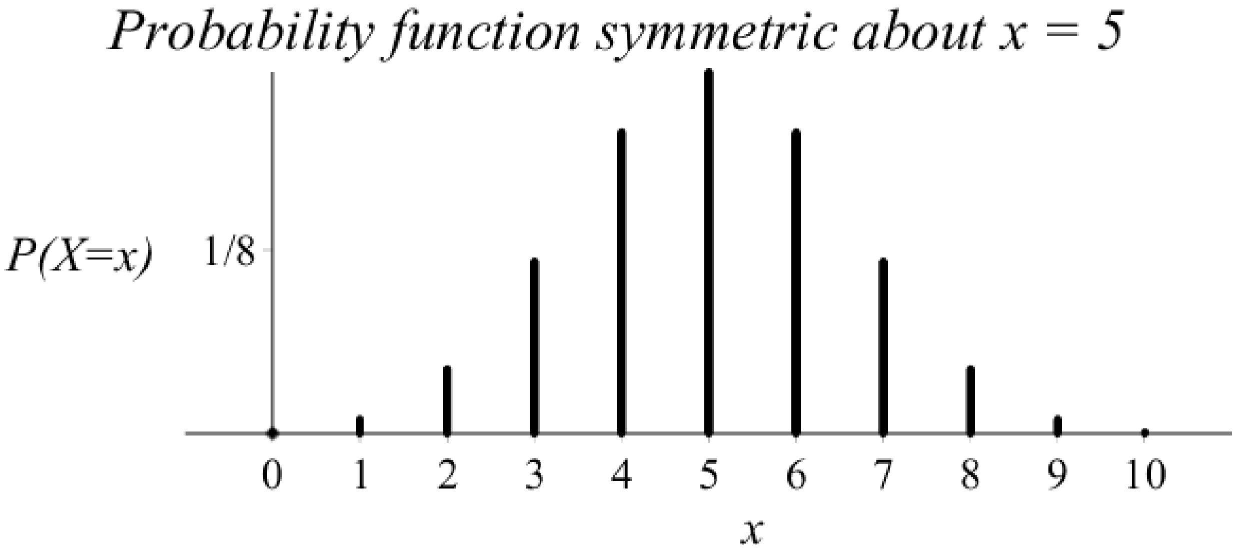
$P(X = x) = \frac{1}{6}$, x = 1,..,6:

$$\boldsymbol{\mu_X = 3.5}$$



Properties of $E(X)$, measure of the centre

1. If the probability function of X is symmetric about $x = c$, then $\boldsymbol{E(X) = c}$



2. $Eg(X) = \sum_{x \text{ in } S_X} g(x)P(X = x)$

3. $E(aX + b) = aE(X) + b$ (a and b: real numbers)

4. $E[ag(X) + bh(X)] = aEg(X) + bEh(X)$

Measure of variation: $var(X)$

The **moments:** $\boldsymbol{E}(\boldsymbol{X}^{k})$ is the $\boldsymbol{k}^{\textbf{th}}$ **moment** van X

$k = 1$: $\mu_X = \boldsymbol{E}(\boldsymbol{X}^{1})$ is the first moment

$k = 2$: $\boldsymbol{E}(\boldsymbol{X}^{2})$ is het 2nd moment: $\boldsymbol{E}(\boldsymbol{X}^{2}) \neq (\boldsymbol{EX})^{2}$

Measures of variation of a r.v. X:

1. The **variance** of X: $\boldsymbol{var}(\boldsymbol{X}) = \boldsymbol{E}(\boldsymbol{X} - \mu_X)^{2}$

Notation: $\boldsymbol{var}(\boldsymbol{X})$ or $\boldsymbol{\sigma}^{2}$ (or σ_X^2)

2. The **standard deviation** $\boldsymbol{\sigma}_{X} = \sqrt{\boldsymbol{var}(\boldsymbol{X})}$

Properties of $var(X)$ and σ_X

1. $\boldsymbol{var(X) \geq 0}$ and $\boldsymbol{\sigma_X \geq 0}$

2. $\boldsymbol{var(X) = E(X^2) - (EX)^2}$

   "the calculation formula"

   Alternative notation: $\boldsymbol{\sigma^2 = E(X^2) - \mu^2}$

3. If $var(X) > 0$, then $E(X^2) > (EX)^2$.

   If $var(X) = 0$, then $P(X = EX) = 1$:

   In the latter case X attains only one value:

   X has a **degenerate distribution.**

4. $var(aX + b) = a^2 var(X)$ and $\sigma_X = |a|\sigma_X$

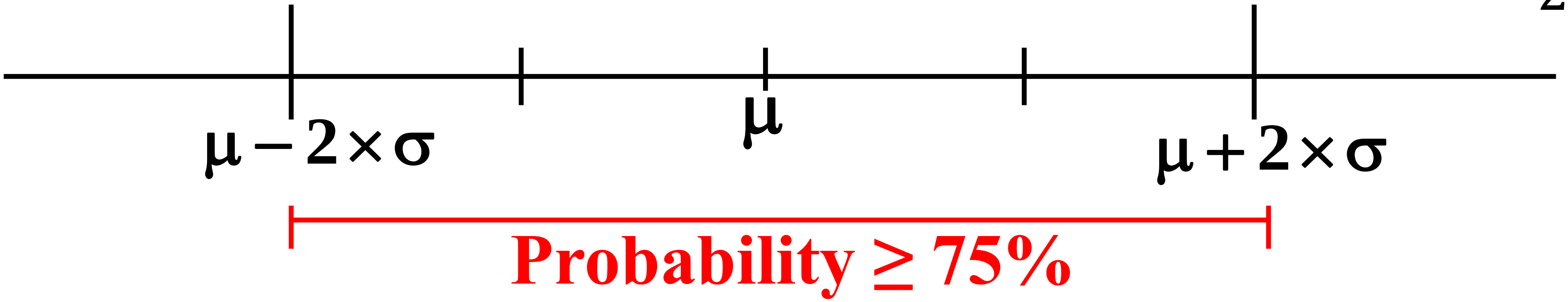
The meaning of the standard deviation σ (1)

For every distribution **Chebyshev's inequality** holds:

$$\textbf{for all c we have: } P(|X - \mu_X| \geq c) \leq \frac{var(X)}{c^2}$$

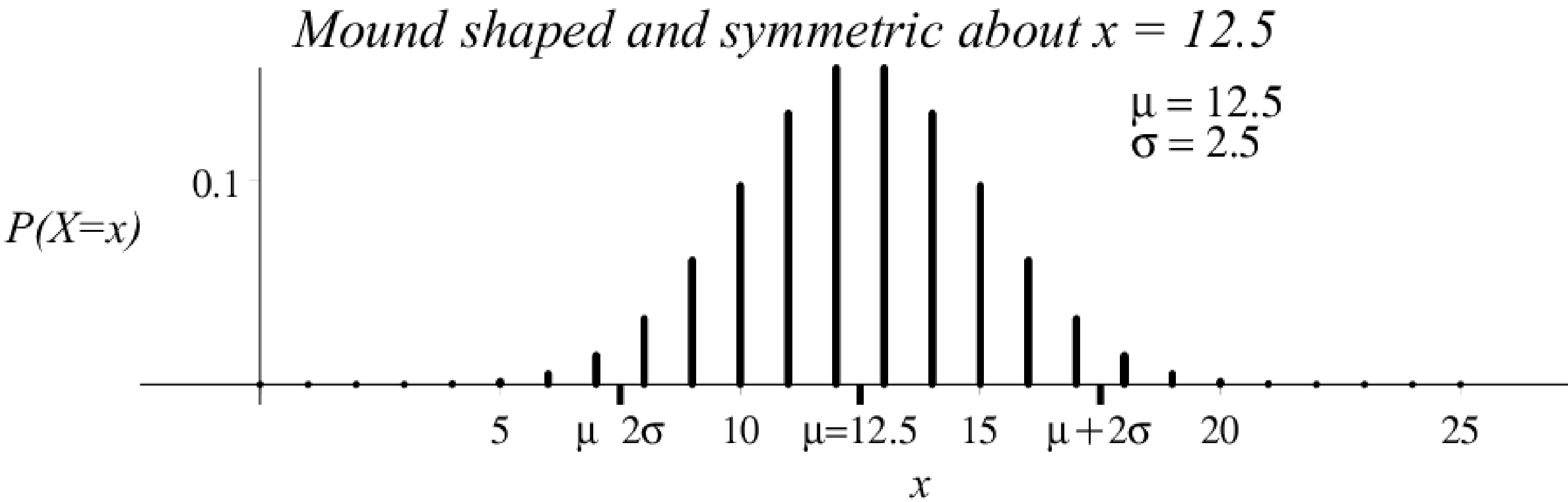
If we choose $c = k \cdot \sigma$ it follows that X attains values in $(\mu - k\sigma, \mu + k\sigma)$ with probability at least $1 - \frac{1}{k^2}$

$\boldsymbol{k = 2}$: X is in $(\mu - 2\sigma, \mu + 2\sigma)$ with prob. $\geq 1 - \frac{1}{2^2}$.



$\boldsymbol{k = 3}$: X is in $(\mu - 3\sigma, \mu + 3\sigma)$ with prob. at least $\frac{8}{9}$

The meaning of the standard deviation σ (2)



The "empirical rule" for σ holds
only for (approximately) "mound shaped" distributions.
Then the following probabilities hold **approximately**:

Interval	***Empirical rule:*** Proportion within interval	***Chebyshev`s rule***
$[\mu - \sigma,\ \mu + \sigma]$	**≈ 68%**	***≥ 0%***
$[\mu - 2\sigma,\ \mu + 2\sigma]$	**≈ 95%**	***≥ 75%***
$[\mu - 3\sigma,\ \mu + 3\sigma]$	**≈ 99.7%**	***≥ 89%***

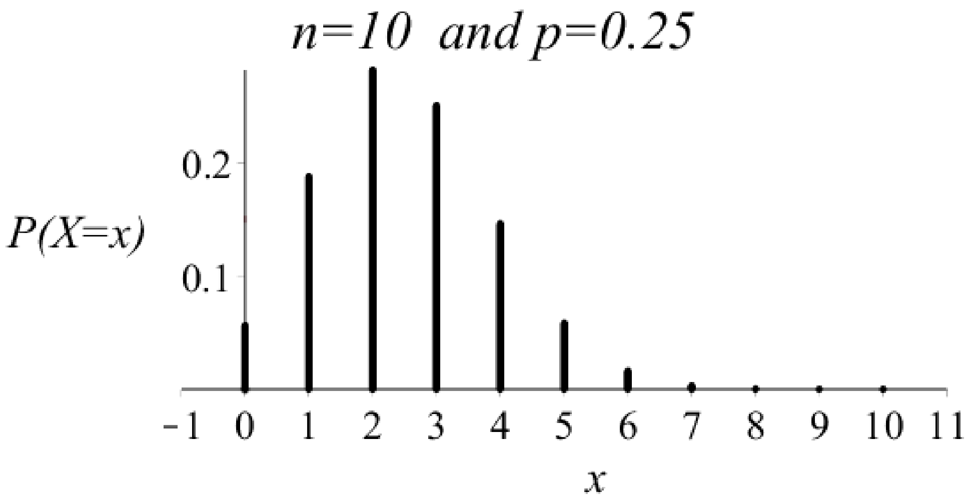
The binomial distribution – 2 examples

Applicable if we have $\boldsymbol{n}$ (ind.) Bernoulli trials with success rate $\boldsymbol{p}$

Example: X = the number of correct random answers to 10 MC-items:

here $\boldsymbol{n} = \mathbf{10}$ and $\boldsymbol{p} = ¼$

$$P(X=x) = \binom{10}{x}\left(\frac{1}{4}\right)^{x}\left(\frac{3}{4}\right)^{10-x},\ x = 0,\ \ldots,10$$



$$E(X) = 10 \cdot \frac{1}{4} = 2.5$$

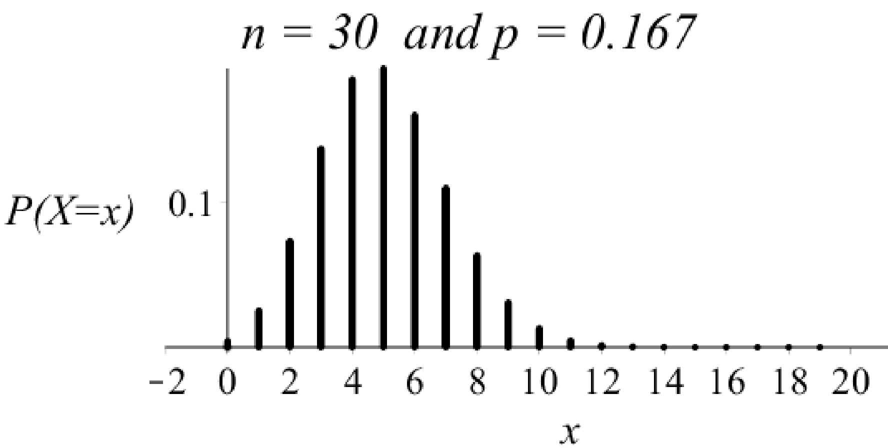
Example:

X = The number of sixes in 30 rolls of a dice.

So $\boldsymbol{n} = \mathbf{30}$ and $\boldsymbol{p} = \frac{1}{6}$

$$P(X=x) = \binom{30}{x}\left(\frac{1}{6}\right)^{x}\left(\frac{5}{6}\right)^{30-x},$$

$$x = 0,\ \ldots,\ 30$$



$$E(X) = 30 \cdot \frac{1}{6} = 5$$

The binomial distribution: $X \sim B(n, p)$

X = "the number of successes in n Bernoulli trials"

Example: count the number of red balls if we draw, at random and with replacement, 8 balls from a box with 10 red and 15 white balls.

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad k = 0, 1, \ldots, n$$

Short notation: X has a $B(n, p)$-distribution

or: $X \sim B(n, p)$

Properties:

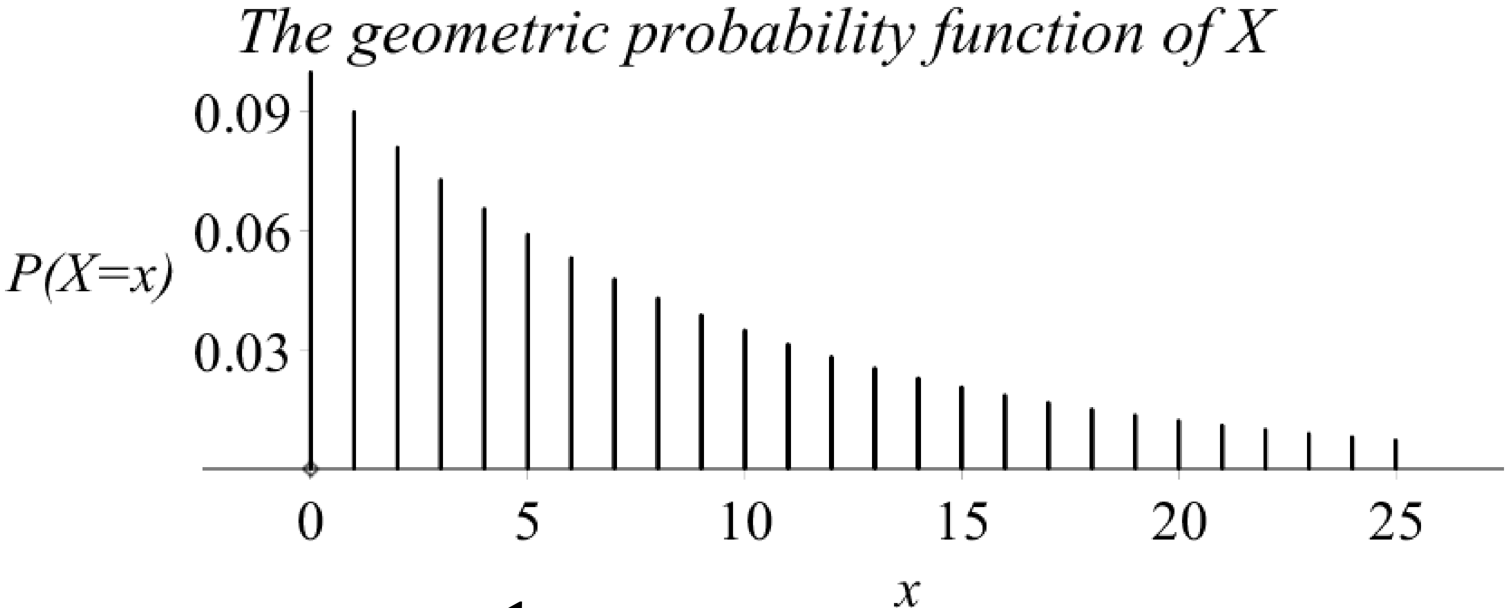
$E(X) = np$ and $var(X) = np(1-p)$

Example geometric distribution

Suppose that 10% of the passing cars are Mercedes.

$X =$ "the number van the first passing Mercedes"

$\boldsymbol{P(X = k) = 0.9^{k-1} \cdot 0.1}$, with **k = 1, 2, 3, …**



We have: $\boldsymbol{E(X)} = \frac{1}{0.1} = \mathbf{10}$

and $P(X > 10) = 0.9^{10} \approx 34.9\%$

The geometric distribution

X = **"number of Bernoulli trials untill the 1st success occurs"**

<u>Example</u>: X is the required number of rolls of a dice to obtain a 6

$$\boldsymbol{P(X=k)=(1-p)^{k-1}p} \quad , \qquad \boldsymbol{k=1,2,3,\ldots}$$

Properties:

1. $\boldsymbol{P(X>k)=(1-p)^{k}}$, for $k = 0, 1, 2, \ldots.$

2. $\boldsymbol{E(X)=\frac{1}{p}}$ and $\boldsymbol{var(X)=\frac{1-p}{p^2}}$

Example hypergeometric formula: the lotto.
6 draws without replacement from balls 1 to 45

Compute the Probability of 4 correct (and 2 wrong)

	correct	wrong	total
Population:	6	39	45
	↓	↓	↓
Sample:	4	2	6

$\boldsymbol{X}$ = "the number of draws with one of my 6 choices"

$$P(X=4)=\frac{\binom{6}{4}\binom{39}{2}}{\binom{45}{6}}\approx 0.14\%$$

and $E(X)=6\times\frac{6}{45}=0.8$

The hypergeometric distribution

X = “the number of red balls if we draw n balls, at random and without replacement, from a box with R red and $N-R$ white balls.

$$\boldsymbol{P(X=k)=\frac{\binom{R}{k}\binom{N-R}{n-k}}{\binom{N}{n}}, k=0,1,\ldots,n}$$

Properties (success probability is $p=\frac{R}{N}$):

1. $\boldsymbol{E(X)=np}$ and $\boldsymbol{var(X)=np(1-p)}$
2. For **large $\boldsymbol{N}$** ($N \geq 5n^2$) $\boldsymbol{X}$ has an **approximate** $\boldsymbol{B(n,p)}$**-distribution** with $p=\frac{R}{N}$

Examples for the Poisson distribution

- ***X* = “the number of car accidents** on a busy junction per week”. *X* varies every week, but from police records we know that the average number is 2.5 per week”:
  ***μ* = 2.5** is the expected number of car accidents per week.
- ***X*, the number of phone calls** at a helpdesk in 10 minutes time interval, can be modelled as a Poisson distributed variables, where ***μ*** is the so called **intensity** of the calls.
  - For half an hour the expected number of calls is 3μ.
- ***X* = “The number of heart attacks on a day** in a town with 100 000 inhabitants”. If the expected number is known, e.g. ***μ* = 2.5**, we can use the Poisson distribution to compute probabilities. But we could also model the number X of heart attacks $B\left(100000, \frac{2.5}{100000}\right)$-distribution.

The Poisson distribution

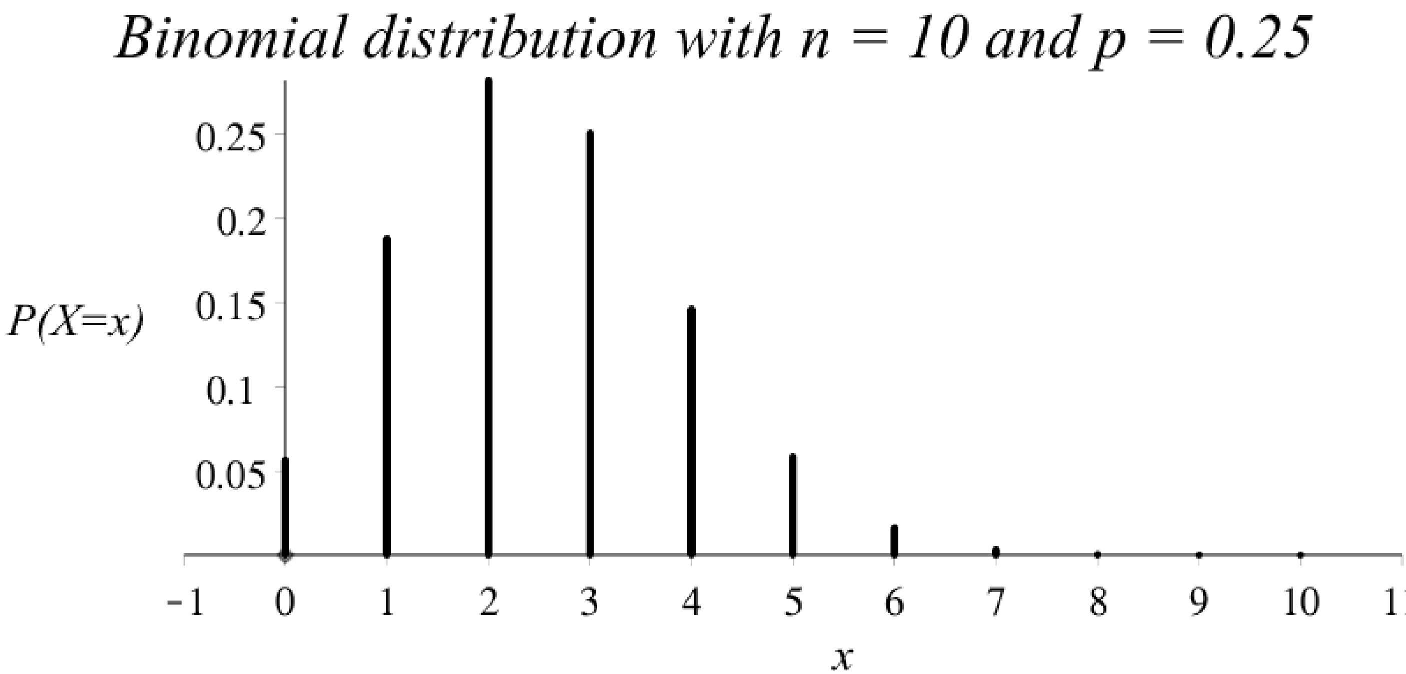
X = "the **number of "rare" events** in a given area and/or interval of time".

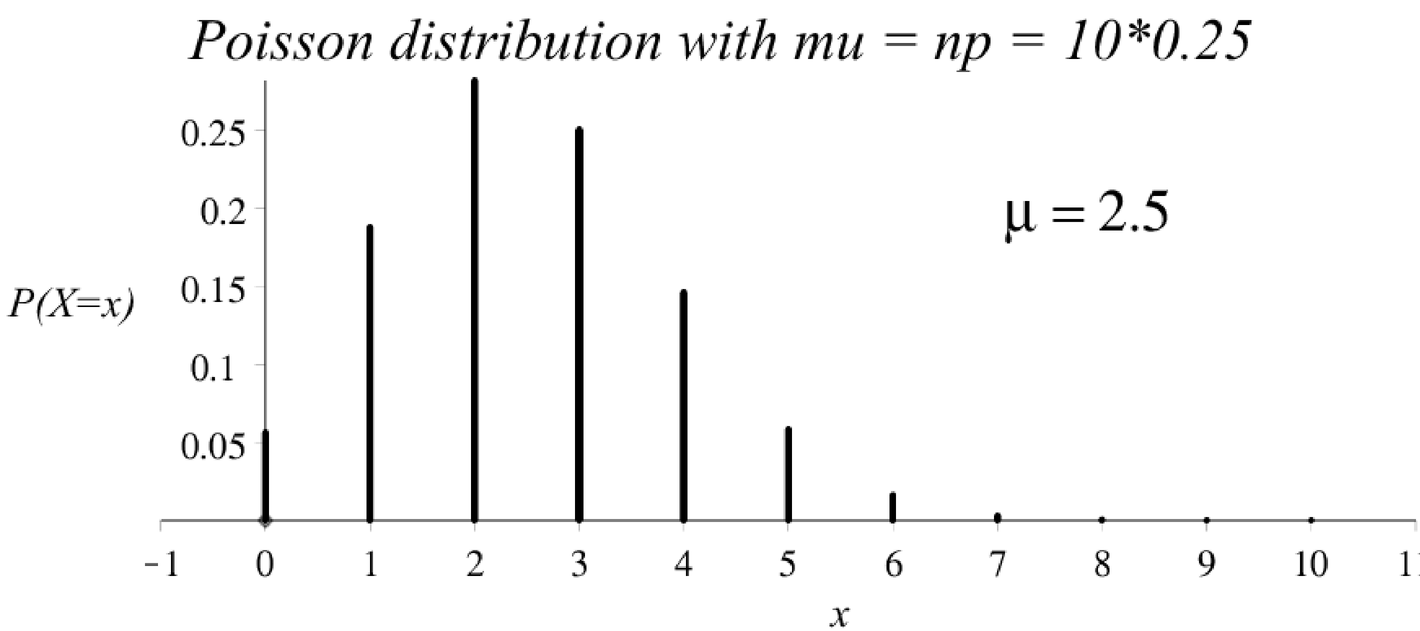
Example: the number of customers entering a system within a 10 min. time interval, on average μ = 2.5

$$P(X = k) = \frac{\mu^k e^{-\mu}}{k!}, \qquad k = 0, 1, \ldots$$

Properties:

1. $\boldsymbol{E(X) = \mu}$ and $\boldsymbol{var(X) = \mu}$
2. If X has a $\boldsymbol{B(n, p)}$**-**distribution with **large $\boldsymbol{n}$ and small $\boldsymbol{p}$** **($\boldsymbol{n \geq 25}$ and $\boldsymbol{np < 10}$ or $\boldsymbol{n(1-p) < 10}$)**, then X is approximately **Poisson** with $\mu = \boldsymbol{np}$**.**


Binomial distribution with n = 10 and p = 0.25
0.25
0.2
0.15
0.1
0.05
P(X=x)
-1
0
1
2
3
4
5
6
7
8
9
10
11
x


Poisson distribution with mu = np = 10*0.25
0.25
0.2
0.15
0.1
0.05
P(X=x)
μ = 2.5
-1
0
1
2
3
4
5
6
7
8
9
10
11
x

Families of discrete distributions

Name + parameter(s)	**Probability function**	$E(X)$	$Var(X)$
Homogeneous on $1, \ldots, n$	$P(X=k) = \frac{1}{n}, k = 1, .., n$	$\frac{n+1}{2}$	**Compute using** $var(X) = E(X^2) - \mu^2$
Binomial $B(n, p)$	$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$, $k = 0, 1, \ldots, n$	np	$np(1-p)$
Geometric (p)	$P(X=k) = (1-p)^{k-1} p$, $k = 1, 2, 3, \ldots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Poisson (μ)	$P(X=k) = \frac{\mu^k e^{-\mu}}{k!}, k = 0, 1, \ldots$	μ	μ
Hyper-geometric (R, N, n)	$P(X=k) = \frac{\binom{R}{k}\binom{N-R}{n-k}}{\binom{N}{n}}$, $k = 0, 1, \ldots, n$	$n \cdot \frac{R}{N}$	$n\frac{R}{N}\left(1-\frac{R}{N}\right) \cdot \frac{N-n}{N-1}$