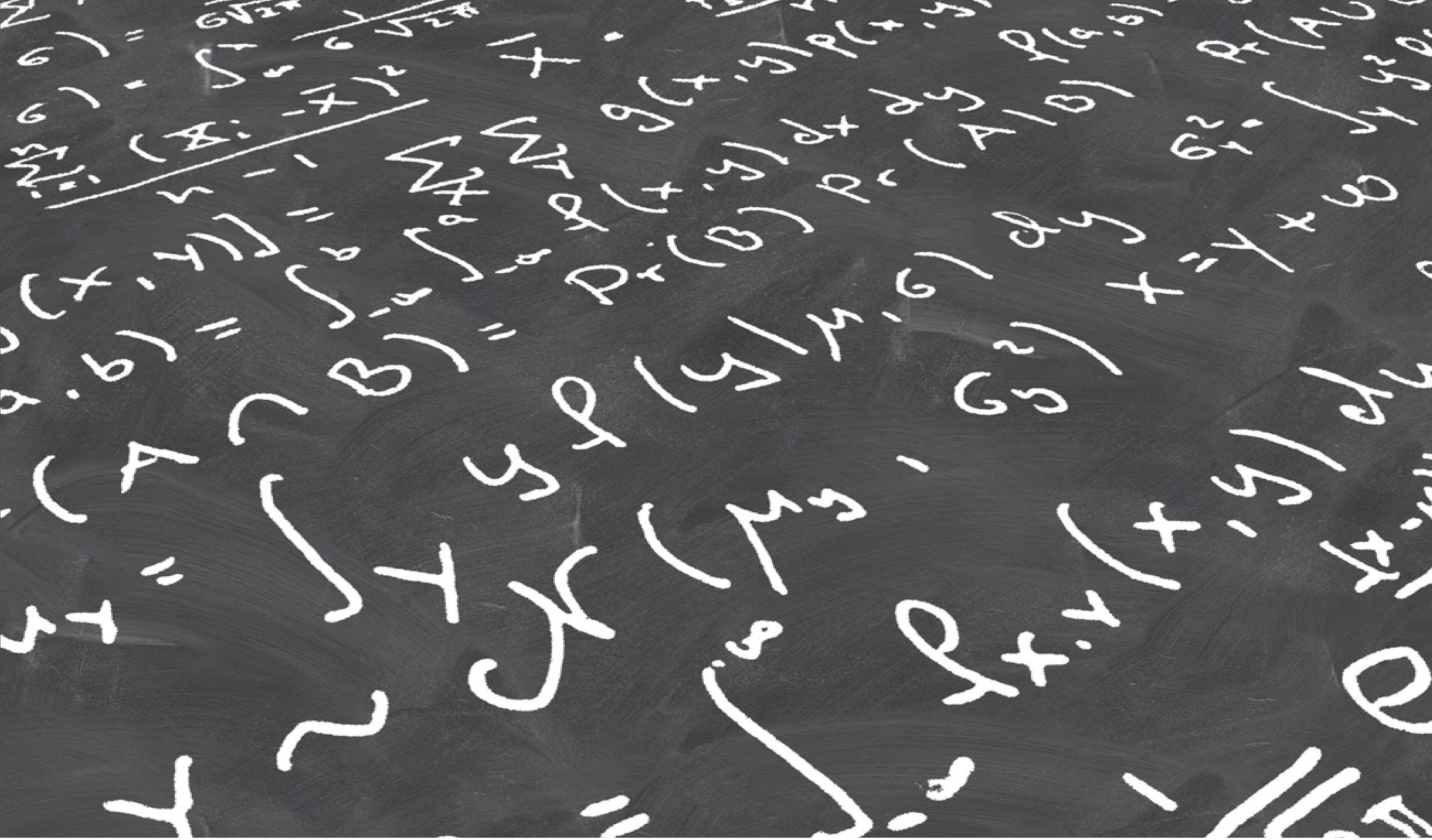


3. Discrete Random Variables

Gunvor Elisabeth Kirkelund
Lars Mandrup

Agenda for Today

- Repetition from last time
- Examples of how information influences probability
- Definition of a Stochastic Random Variable
- Discrete Stochastic Variables
- Discrete Stochastic Distributions
- Mean, Variance and Standard Deviation
- Some common Discrete Probability Distributions

Bayes Rule and Independence

- Bayes Rule:

$$\boxed{Pr(A|B) = \frac{Pr(A \cap B)}{Pr(B)} = \frac{Pr(B|A) \cdot Pr(A)}{Pr(B)}}$$

$$\boxed{Pr(A \cap B) = Pr(A|B) \cdot Pr(B) = Pr(B|A) \cdot Pr(A)}$$

- A and B independent:

$$\boxed{Pr(A \cap B) = Pr(A) \cdot Pr(B)}$$

$$\boxed{Pr(B|A) = Pr(B) \quad \text{and} \quad Pr(A|B) = Pr(A)}$$

Combinatorics

- The number of possible outcomes of k trials, sampled from a set of n objects.

Types of Experiments:

- With or without replacement
- Ordered or unordered

		Replacement	
		With	Without
Sam-pling	Ordered	n^k	$P_k^n = \frac{n!}{(n-k)!}$
	Unordered	$\binom{n+k-1}{k} = \frac{(n+k-1)!}{k!\,(n-1)!}$	$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$

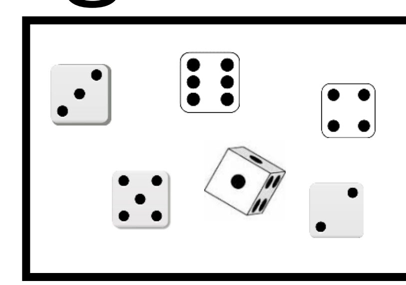
Also called a random experiment

Stochastic Experiment

- An experiment in which you can not predict the outcome

Examples:

S



- Rolling a dice
- Sample space for the experiment is: {1, 2, 3, 4, 5, 6}

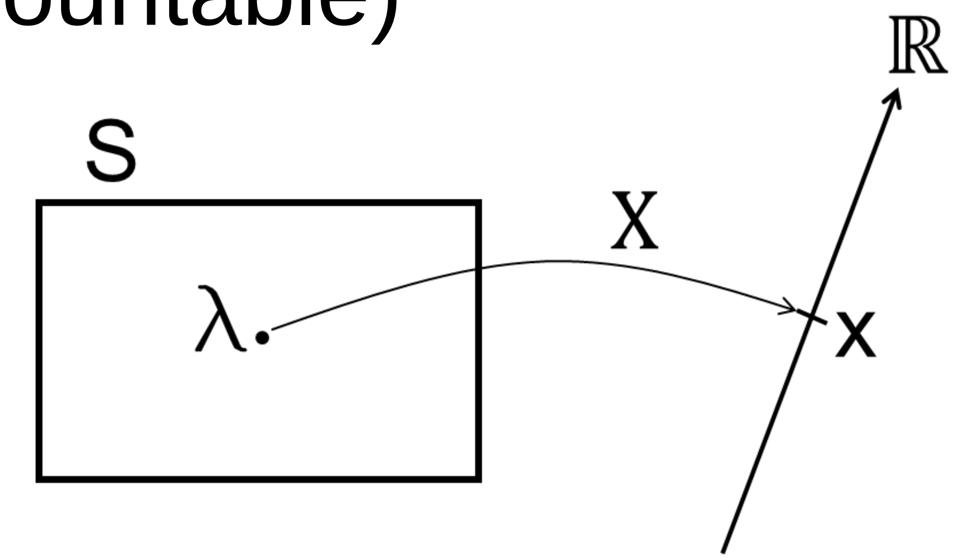
S

- Flip a coin
- Sample space for the experiment is: {Head, Tail}

Also just called a random variables

Stochastic Random Variables

- A random variable tells something important about a stochastic experiment.
- Can be discrete (R_X= range of X, countable) or continuous (R_X = range of X, uncountable)



Examples:

- The numbers on a dice (discrete):
  - Sample space for variable X is: {1, 2, 3, 4, 5, 6}
  - Sample space for variable Y “Even (1)/Uneven (-1)”: {1, -1}
- The hight of students at ECE (continous):
  - Sample space for variable H is all real numbers: [100;250] cm.

Probability Mass Function (PMF)

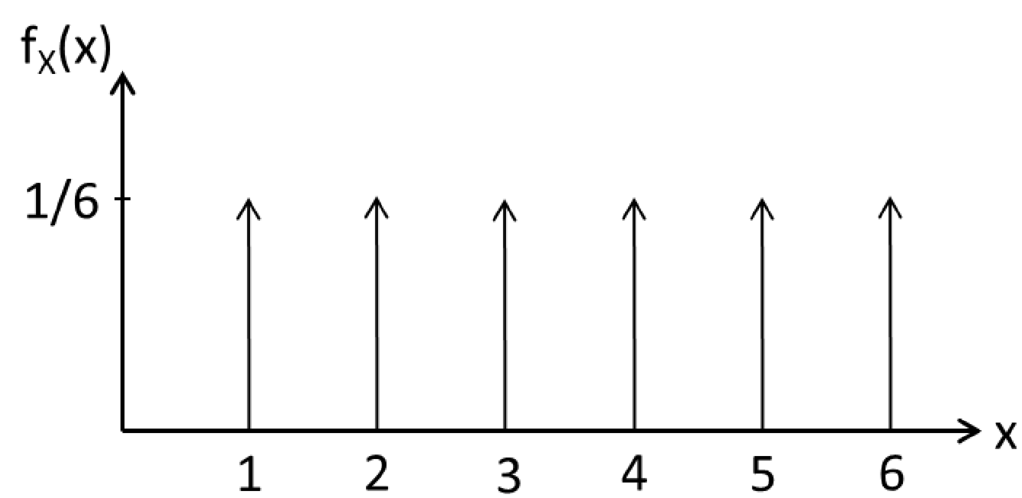
- Range of X, R_X, countable.
- X is a <u>discrete</u> stochastic variable.

$$f_X(x) = \boxed{\begin{cases} Pr(X = x_i) & for\ X = x_i \\ 0 & otherwise \end{cases}}$$

$$0 \leq f_X(x) \leq 1$$

- We have that: $\displaystyle\sum_{x_i \in R_X} f_X(x_i) = \sum_{x_i \in R_X} Pr(X = x_i) = 1$

Example: Perfect dice



Cumulative Distribution Function (CDF)

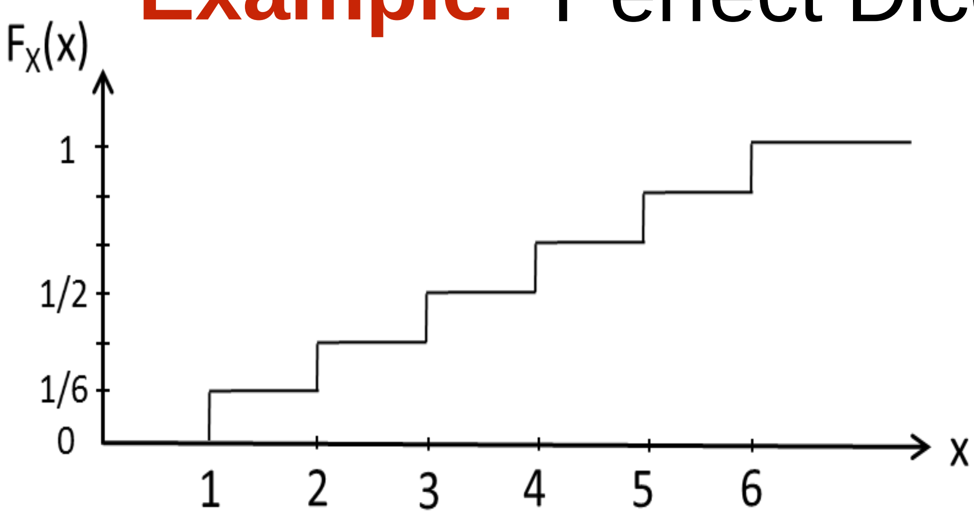
- Range of X, R_X, countable.
- X is a <u>discrete</u> stochastic variable.
- $F_X(x)$ is a non-decreasing staircase-function.

$$\boxed{F_X(x) = Pr(X \leq x) = \sum_{x_i=-\infty}^{x_i \leq x} f_X(x_i) = \sum_{x_i \in R_X} f_X(x_i) u(x_i - x)}$$

We have that:

- $0 \leq F_X(x) \leq 1$
- $\lim_{x \to -\infty} F_X(x) = 0, \quad \lim_{x \to \infty} F_X(x) = 1$
- Steps $= Pr(X = x_i) = f_X(x_i)$
- $F_X(x_2) - F_X(x_1) = Pr(x_1 < X \leq x_2)$

Example: Perfect Dice



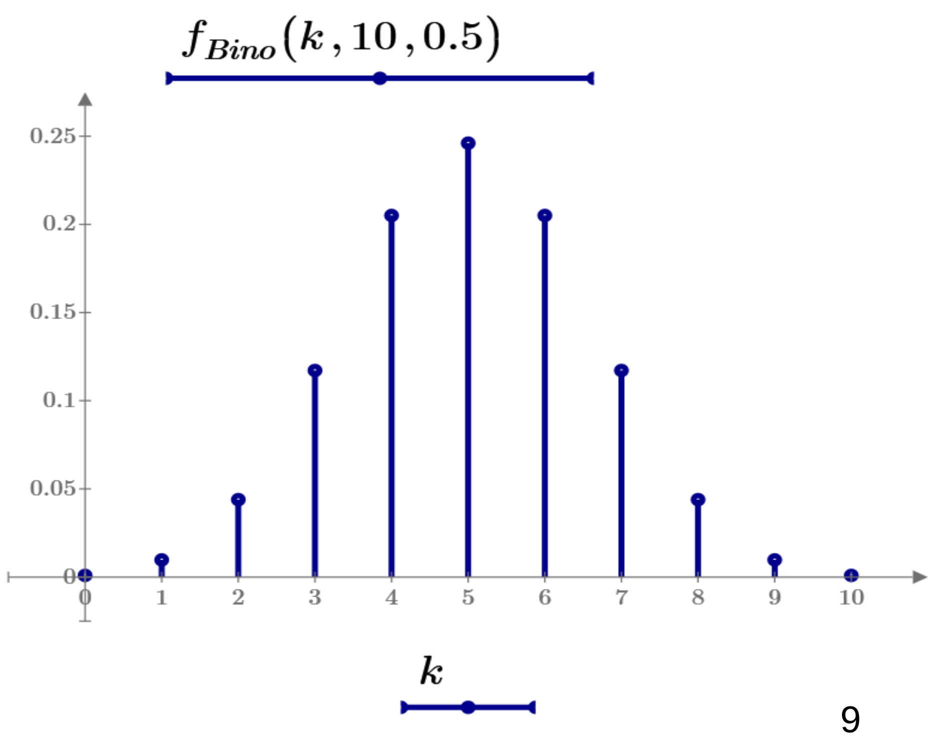
The Binomial Mass Function

- We have n repeated trials.
- Each trial has two possible outcomes
  - **Success** — probability p
  - **Failure** — probability 1-p

Also called a Bernoulli trial

- X is the number of successes k in n trials
- $X \sim Binomial(n, p)$
- The probability mass function for X is given as:

$$\boxed{f(k|n,p) = \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k}}$$



The Binomial Distribution

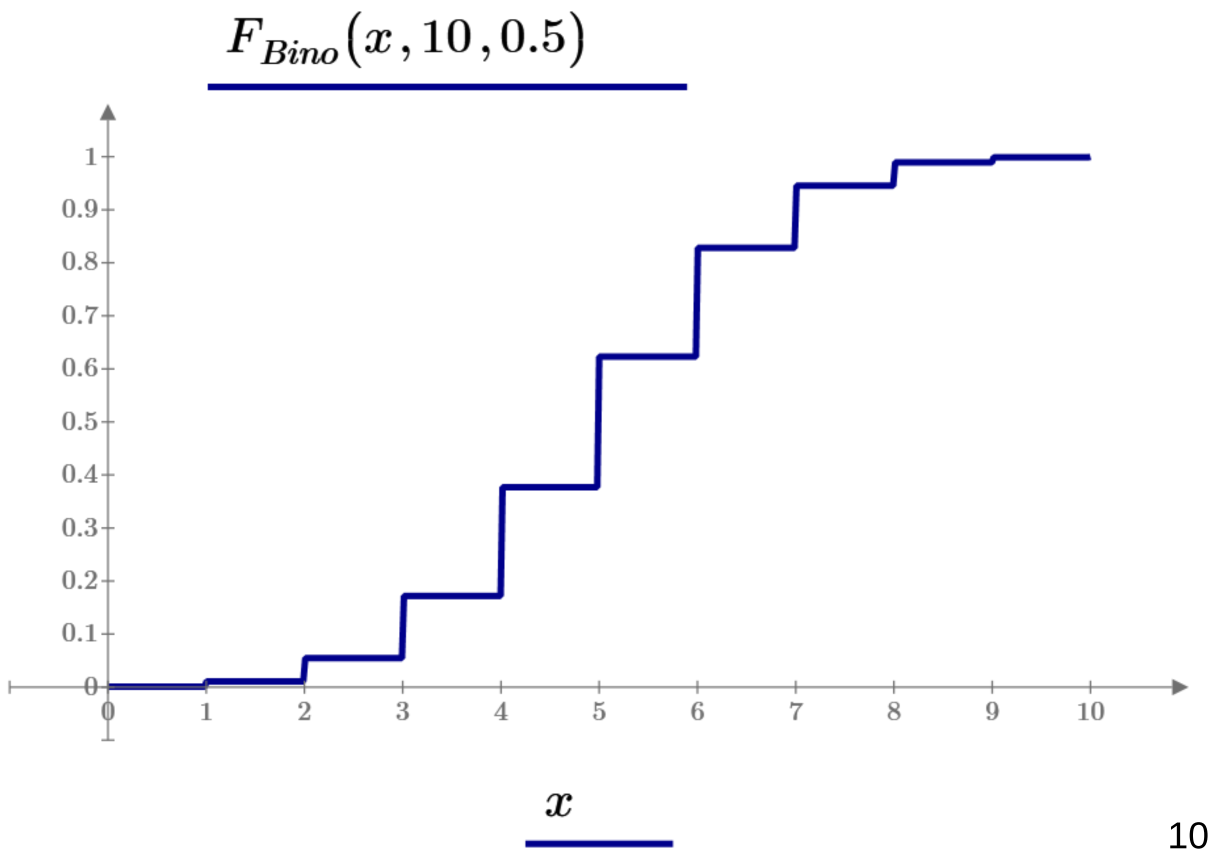
- $X \sim Binomial(n, p)$

- The probability mass function is given as:

$$f(k|n,p) = \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}$$

- We write the distribution as the sum:

$$F(k|n,p) = \sum_{i=0}^{k} f(i|n,p)$$



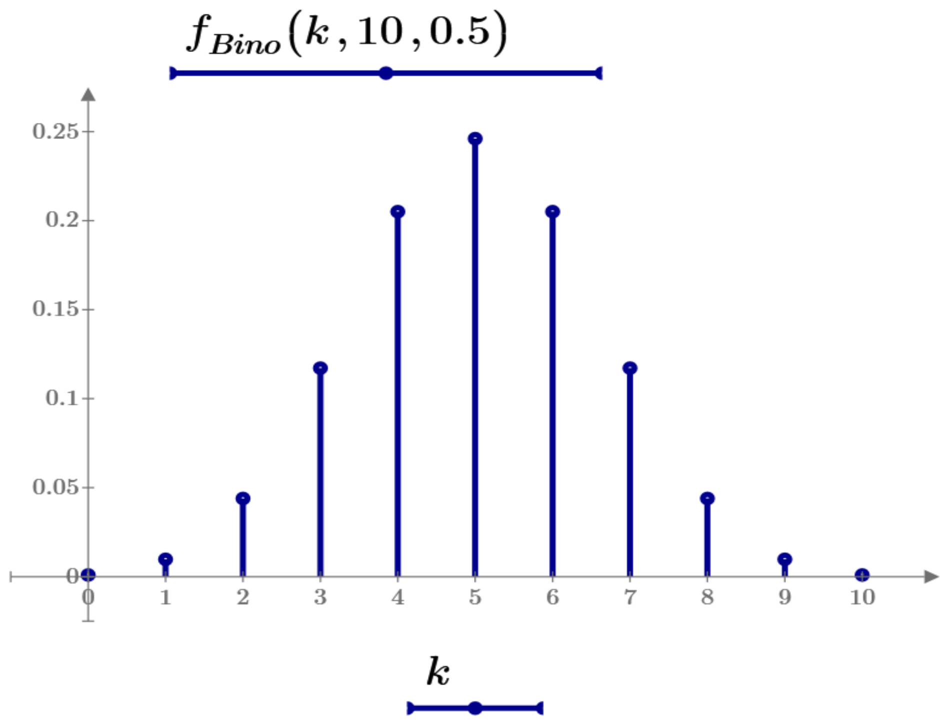
Expectation of a Discrete Random Variable

Example: If I want ten children, how many girls can I expect to get?

Answer: I assume a Binomial distribution with p=0.5:

$$f(k|10{,}0.5) = \binom{10}{k} \cdot 0.5^k \cdot 0.5^{10-k} = \binom{10}{k} \cdot 0.5^{10}$$

where: $\binom{10}{k} = \frac{10!}{k!(10-k)!}$



$$E[k] = 0 \cdot f(0|10{,}0.5) + 1 \cdot f(1|10{,}0.5) + \cdots + 10 \cdot f(10|10{,}0.5)$$

$$= \left(0 + 1 \cdot \binom{10}{1} + 2 \cdot \binom{10}{2} \ldots + 10 \cdot \binom{10}{10}\right) \cdot 0.5^{10}$$

$$= (0 + 1 \cdot 10 + 2 \cdot 45 + \cdots + 10 \cdot 1) \cdot 0.5^{10} = 10 \cdot 0.5 = 5$$

Expectation of a Discrete Random Variable

- We define the mean or the expectation of a discrete random variable as:

n is the number of outcomes

$$\boxed{EX = E[X] = \bar{X} = \mu_X = \sum_{i=1}^{n} x_i \, f_X(x_i)}$$

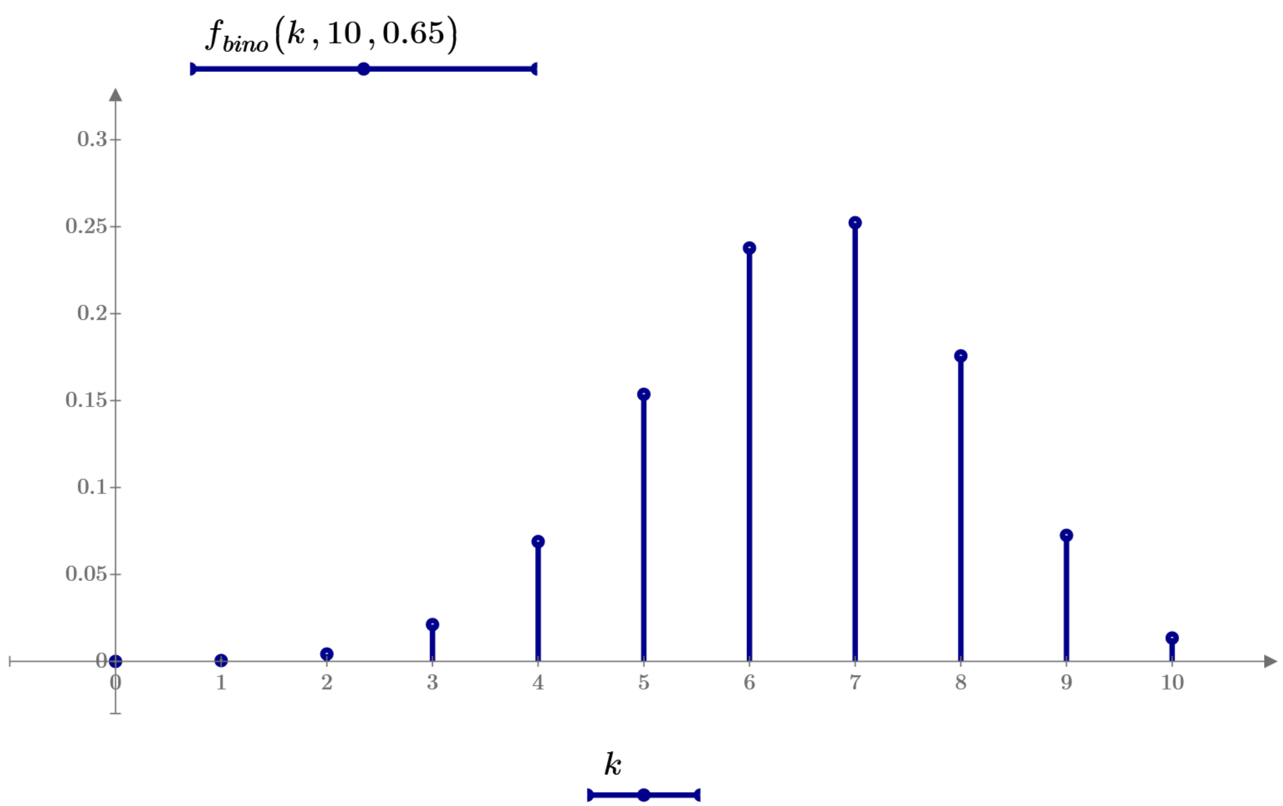
$f_X(x_i)$ is the probability (frequency) of outcome x_i

x_i is its outcome

The Binomial Distribution (cont'd)

- For the Binomial distribution, we have:

$$E[k] = n \cdot p$$
$$Var(X) = n \cdot p \cdot (1 - p)$$



- Where the variance is defined as:

$$Var(X) = \sigma_X^2 = E\left[X^2\right] - E[X]^2$$

Variance and standard deviation tells of the spreading of the data

Variance and standard deviation

- The variance σ^2 is an indicator on how much the values of a random variable X are spread around (deviates from) the expectation value:

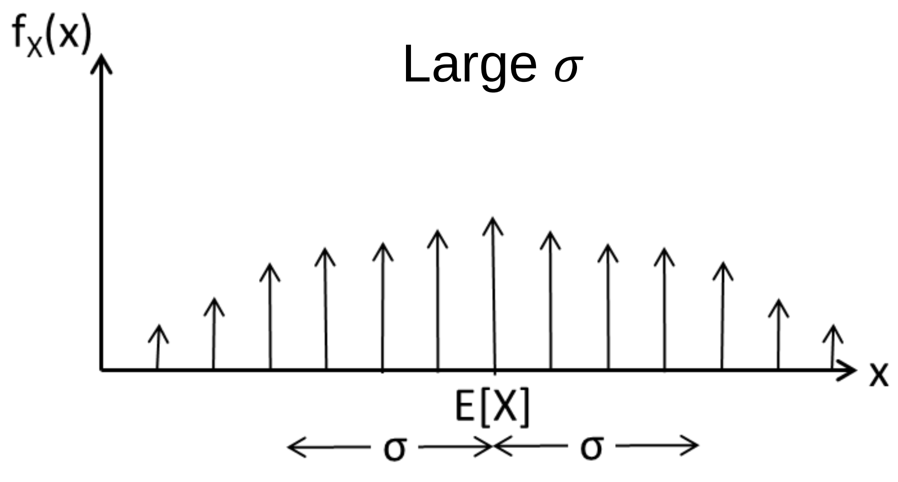
$\sum_x x^2 f_X(x)$

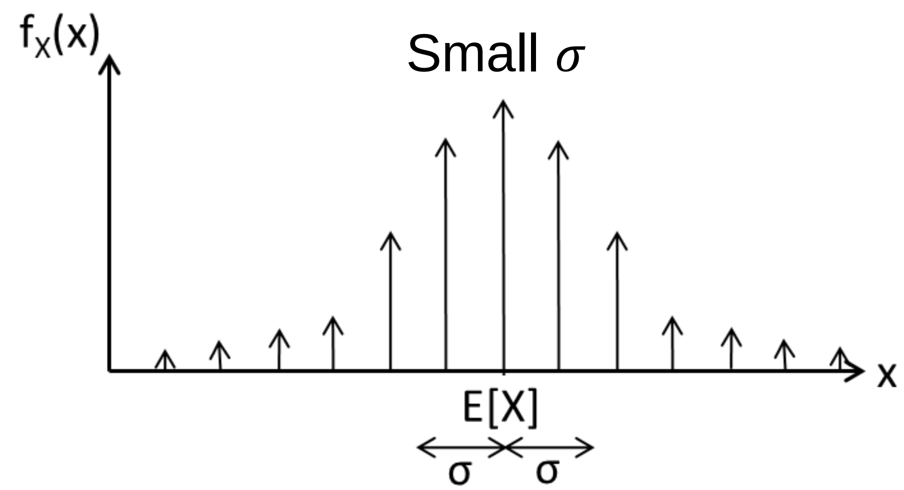
$$\boxed{Var(X) = \sigma_X^2 = E\left[(X - \mu_X)^2\right] = E\left[X^2\right] - E[X]^2}$$

$\left(\sum_x x f_X(x)\right)^2$

- The standard deviation σ is the square root of the variance (same unit as X):

$$\boxed{SD(X) = \sigma_X = \sqrt{\sigma_X^2}}$$





Functions of random variables

- $\mathrm{Y} = g(X)$:

  - $f_Y(y) = Pr(Y = y) = Pr(g(X) = \mathrm{y})$

    $= \sum_{x:\, g(x)=y} Pr(X = x) = \sum_{x:\, g(x)=y} f_X(x)$

  - $EY = E[Y] = E[g(X)] = \overline{g(X)} = \sum_x g(x) f_X(x)$

    LOTUS – Law Of The Unconscious Statistician

  - $Var(Y) = Var(g(X)) = E\left[g(X)^2\right] - E[g(X)]^2$

Expectation and Variance of linear combinations

- Linear function: $\mathrm{Z} = g(X) = aX + b$
  - $EZ = E[Z] = \mu_Z = E[aX + b] = \sum_x (ax + b) \cdot f_X(x) = a \cdot EX + b$
  - $Var(Z) = \sigma_Z^2 = Var(aX + b) = E\big[(aX + b)^2\big] - E[aX + b]^2$
    $= a^2 \cdot \big(E[X^2] - E[X]^2\big) = a^2 \cdot Var(X)$
  - $SD(Z) = \sigma_Z = |a| \cdot \sigma_X$

- Linear sum: $\mathrm{Z} = aX + bY$ (X and Y independent)
  - $EZ = E[Z] = \mu_Z = E[aX + bY] = a \cdot EX + b \cdot EY$
  - $Var(Z) = \sigma_Z^2 = Var(aX + bY) = a^2 \cdot Var(X) + b^2 \cdot Var(Y)$
  - $SD(Z) = \sigma_Z = \sqrt{a^2 \cdot \sigma_X^2 + b^2 \cdot \sigma_Y^2} \neq |a| \cdot \sigma_X + |b| \cdot \sigma_Y$

OBS!

Discrete Uniform Distribution - pmf

- $X \sim \mathcal{U}[a, b]; \quad a, b \in \mathbb{Z}$
- All integer numbers between a and b have equal probability

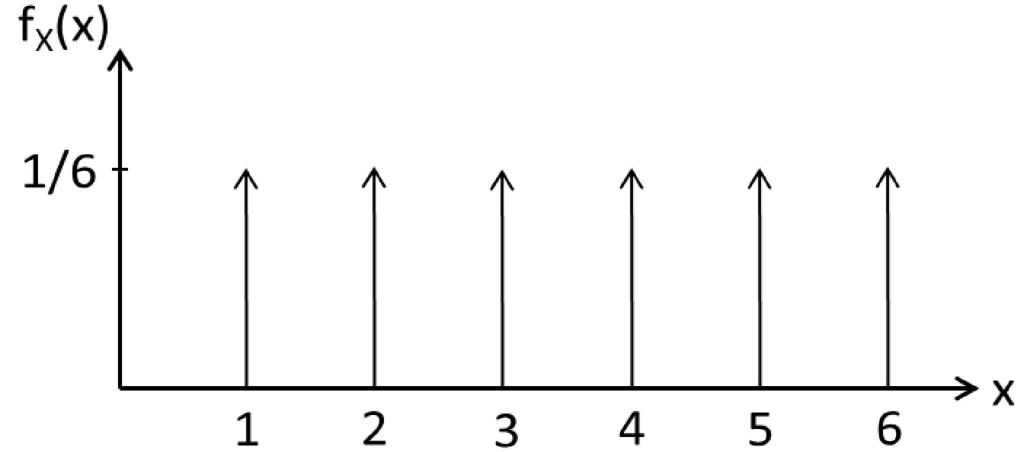
$$\boxed{f_X(x) = \begin{cases} \dfrac{1}{b - a + 1} & for\ x = a, a + 1, \ldots, b \\ 0 & otherwise \end{cases}}$$

$$\sum_{x=a}^{b} f_X(k) = \sum_{x=a}^{b} \frac{1}{b - a + 1} = 1$$

$$EX = \frac{a + b}{2} \qquad Var(X) = \frac{(b - a + 1)^2 - 1}{12}$$

Example: Perfect dice

$$f_X(x) = \begin{cases} \frac{1}{6} & for\ 1 \le x \le 6 \\ 0 & otherwise \end{cases}$$



Discrete Uniform Distribution - cdf

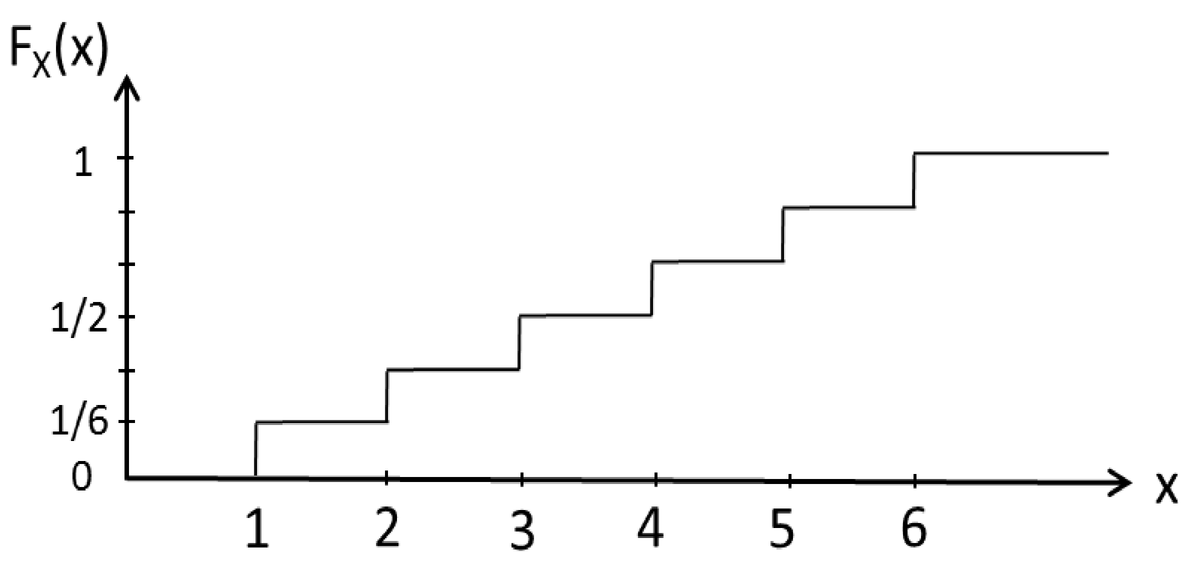
- $X \sim \mathcal{U}[a,b]; \quad a,b \in \mathbb{Z}$
- All integer numbers between a and b have equal probability
- $F_X(x)$ a staircase-function with equal step-size

$$\boxed{F_X(x) = Pr(X \leq x) = \sum_{x_i=a}^{x_i \leq x} \frac{1}{b-a+1} = \frac{1}{b-a+1} \sum_{x_i=a}^{b} u(x_i - x)}$$

$$0 \leq F_X(x) \leq 1$$

Example: Perfect dice

$$F_X(x) = \begin{cases} 0 & for \quad x < 1 \\ \frac{i}{6} & for\; i \leq x < i+1 \\ 1 & for \quad x \geq 6 \end{cases}$$



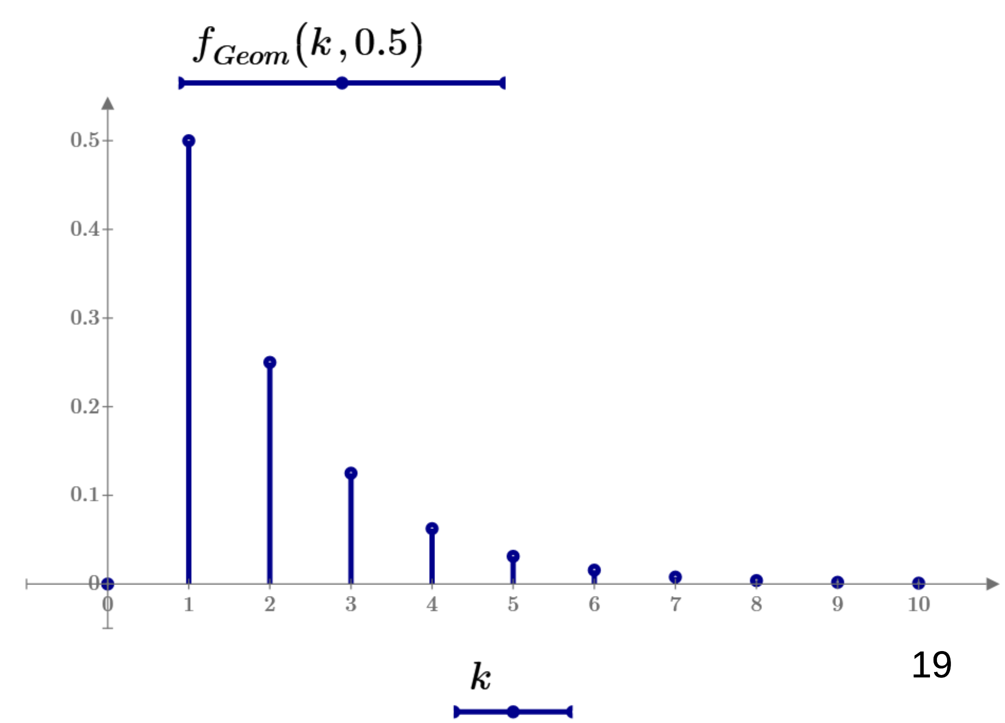
Geometric Distribution - pmf

- Repeated Bernoulli trial $B(p)$
- X the total number of trials until the first succes
- $X \sim Geometric(p)$

$$f_X(k) = \begin{cases} p(1-p)^{k-1} & for\ k = 1,2,3, \dots \\ 0 & otherwise \end{cases}$$

$$EX = \frac{1}{p}$$

$$Var(X) = \frac{1-p}{p^2}$$



Examples:

- Flip a coin
- X = Number of flips until the first Head

Poisson Distribution - pmf

- X the number of events k occurring in a fixed interval of time t, when:
  - these events occur with a known average rate λ
  - the events are independent of the time since the last event
- $X \sim poisson(\lambda)$

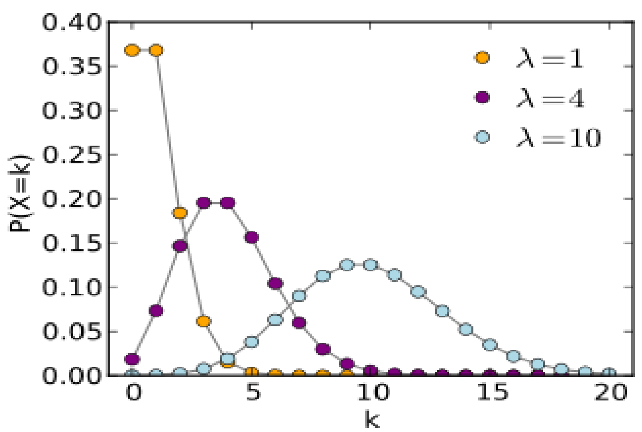
$$f_X(k) = \begin{cases} \dfrac{\lambda^k e^{-\lambda}}{k!} & for\ k = 0,1,2,3,\ldots \\ 0 & otherwise \end{cases}$$

$$EX = \lambda$$

$$Var(X) = \lambda$$

Examples:

- Number of customers in 1 hour



Words and Concepts to Know

Stochastic

Expectation

Cumulative Distribution Function

Probability Mass Function

Geometric Distribution

Uniform Distribution

Discrete stochastic variable

Expectation value

Staircase-Function

Standard deviation

Binomial Mass Function

Mean

pmf

Variance

Poisson Distribution

Binomial Distribution

cdf