Counting

Chapter 3

Random variables

rastgele deneyin olası sonuçlarına sayısal değer atayan gerçek değerli bir fonksiyondur. random variable aslında bir fonksiyondur

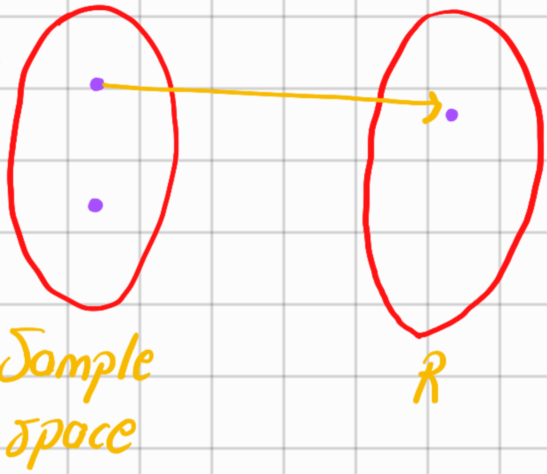
Random variable is a real-valued function that assigns numerical values to possible outcomes of the random exp.

Büyük harfle gösterilir

A random variable X is a function from sample space to real numbers

$$X : S \longrightarrow R$$

random variable bir elementi örnek uzaydan bir sayıya eşler



Example : toss a coin 3 times

X is the number of heads

range of X ?

rastgele deneyler için büyük harf kullanıyoruz

Range (x) R_X $\{0, 1, 2, 3\}$

Range (X) = $R_X = \{0,1,2,3\}$

→ Bunun anlamı şudur: X'in olası değerlerinin kümesi demek

→ Bu random variable'ın range'idir.

HHH → 3
HHT → 2
HTT → 1
TTT → 0
THH → 2
TTH → 1
THT → 1
HTH → 2

→ Bu küme bize aralıktaki elemanları gösterir

→ 2^3

Sorudaki X yazı tura 3 kez atıldığındaki tura sayısıdır

random experimentları göstermek için biz büyük harf kullanırız

random variable sadece bir fonksiyondur

aslında örnek uzaydaki her olası sonuca sayısal bir değer atar

Example:

① Toss a coin 100 times. X is the number of tails?

$R_X = \{0, 1, 2, \dots, 100\}$

Bunun anlamı X'in range'i demek

0 → hiç yazı gelmemesi

100 → hepsinin yazı olması

→ Countable, finite

② I toss a coin until the first tail. Let Y be the number of tosses

$R_Y = ?$ (range of Y demek)

$\rightarrow$ H H H H T

$\rightarrow$ T

$\rightarrow$ H ---- H T

$R_y = \{1, 2, \text{---}\} = \mathbb{N}$

T HT infinity

Countable infinite

if the range R is countable, X is a discrete random variable

$\mathbb{Q}, \mathbb{N}, \mathbb{Z} \} \rightarrow$ countable

$\mathbb{R} \rightarrow$ Reel numbers $\rightarrow$ uncountable

eger range uncountable ise biz buna Continuous random variable deriz

---

Probability mass function

Let X be a random variable with range

$R_X = \{x_1, x_2, \text{---}\}$ (finite, countably infinite)

The function

The function

$P_X(x_k) = P(X = x_k)$ for $k = 1, 2, 3 \dots$ sonsuza kadar gider

is called probability mass function of X

PMF

---

example: toss a coin twice, X is number of heads

1_ Range of X? $R_X = \{0, 1, 2\}$

2_ P_X?

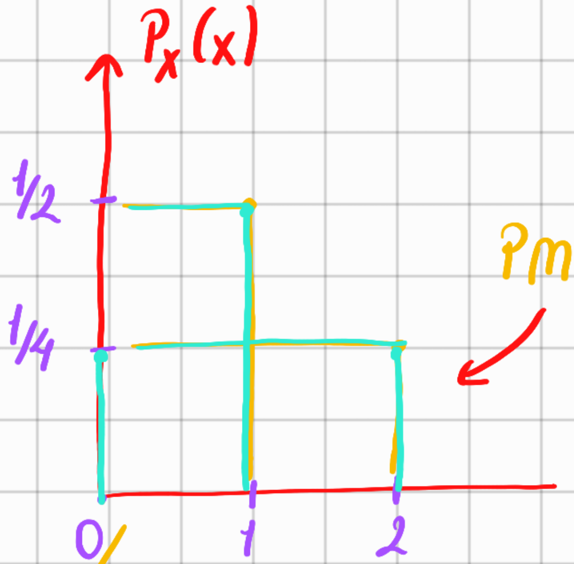
HH	2
HT	1
TH	1
TT	0

$P_X(0) = (X=0) = 1/4$

$P_X(1) = P(X=1) = 2/4 = 1/4 + 1/4$

$P_X(2) = P(X=2) = 1/4$

the distribution function is usually reserved for the Cumulative distribution



PMF plot

probability distribution

Cumulative distribution function

Example: We have unfair coin $P(H) = p \quad 0 < p < 1$

I toss the coin until I get a H. Y is the total number of tosses. Distribution of Y?

$$R_Y = \mathbb{N} = \{1, 2, 3, \ldots\}$$

$$P_Y(1) = (Y=1) = p$$

$$TH \leftarrow P_Y(2) = (Y=2) = (1-p).p$$

$(1-p)$

$$TTH \leftarrow P_Y(3) = (Y=3) = (1-p)(1-p).p$$

$k^{th} \to H$

$$P_Y(k) = (Y=k) = (1-p)^{k-1}.p$$

$(k-1)$

$\underbrace{TTTT\_\_\_T}_{k-1}$

$$P_Y(k) = \begin{cases} (1-p)^{k-1} \cdot p & \text{for } k = 1,2,3\ldots \\ 0 & \text{otherwise} \end{cases}$$

→ probability mass function

PMF is a probability measure

1 → $0 \leq P_X(x) \leq 1$ for all x

2 → $\sum_{x \in R_X} P_X(x) = 1$

3 → for any set $A \subset R_X$

$P(X \in A) = \sum_{X \subset A} P_X(x)$

olasılık hakkındaki 3 axioms

→ probability mass function $P_X(x)$ tüm x'ler için 0 ve 1 arasında

→ Bu kuralların hepsi PMF için geçerlidir

Example: Let X be discrete random variable

$R_X = \{1,2,3,4\}$, $P_X(1) = P_X(2) = 1/3$ and $P_X(4) = 1/6$

a-) $P_X(3) = ?$ $\sum_{x \in R_X} P_X(x) = 1 = 1/3 + 1/3 + P_X(3) + 1/6 = 1$

→ $P_X(3) = 1/6$

b-) $P(2 \leq X < 4) = ?$

→ $P(X \in \{2,3\})$

$= P_X(2) + P_X(3)$

$= 1/3 + 1/6 = 1/2$

$\frac{1}{3} + \frac{1}{6} = \frac{1}{2}$

Independent random variables

(x ve y birbirini etkilemiyorsa x ve y independent olur)

Previous independent events

$P(A \cap B) = P(A) \cdot P(B)$ Bunu sağlıyorsa independent eventtir

Consider two discrete random variables X and Y.

X and Y are independent if

$P(X=x, Y=y) = P(X=x) \cdot P(Y=y)$ Bu olay için 2 tane discrete random variables gerekiyor

$P(X \in A, Y \in B) = P(X \in A) \cdot P(Y \in B)$

Aslında 2 tane random variable yani X ve Y birinin değeri diğerinin olasılığını değiştirmiyorsa independent olur

example: Roll a die twice and get two numbers X and Y

a-) find R_x, R_y and PMF of X and Y

$R_x = R_y = \{1,2,3,4,5,6\}$

$P_X(1) = 1/6 = P_X(2) = P_X(3) = P_X(4) = P_X(5) = P_X(6)$

$P_Y(1) = 1/6 = P_Y(2) = P_Y(3) = P_Y(4) = P_Y(5) = P_Y(6)$

b) find $P(x=2 \text{ and } y=6) = P_X(2) \cdot P_Y(6) = 1/6 \cdot 1/6$

(Bu independent random variable'a bir örnek)

$\hookrightarrow 1/36$

ⓒ find $P(\underbrace{x>3}_{4,5,6} \text{ and } Y=2) = \underbrace{[P_X(4)+P_X(5)+P_X(6)]}_{1/2 \cdot 1/6 \to 1/12} \cdot P_Y(2)$

ⓓ Let $Z = X+Y$ Range of $Z=?$

$Z = X+Y$

$R_X = R_Y = \{1,2,3,4,5,6\}$

$P_X(1) = P_X(2) \ldots P_X(6) = 1/6$

$$P_Z(k) = \begin{cases} 1/6 & k=\{1,2 \ldots 6\} \\ 0 & \text{otherwise} \end{cases}$$

$P_Z(2) = P(X+Y=2)$

$P(X=1, Y=1) = 1/6 \cdot 1/6 = 1/36$

$P_Z(3) = P(X+Y=3) = P(X=1, Y=2)$ or $P(X=2, Y=1)$

$= P(X=1) \cdot P(y=2) + P(X=2)\, P(Y=1)$

$P_Z(4) = P(X+Y=4) \to X=1 \quad y=3$ ... Bu ...

$P_Z(4) = P(X+Y=4)$ → $X=2 \quad y=2$ + / $X=3 \quad y=1$ +

bunların olasılık-larını topla

Example: I toss a coin twice
X = number of heads I observe
I toss the con two more times
Y number of heads I observe

$P(X<2 \text{ and } Y>1) = ?$ → independent olaylar

Solution = $P((X<2) \text{ and } (Y>1)) = P(X<2) \cdot P(Y>1)$

$\left[P_X(0) + P_X(1)\right] \cdot P_Y(2)$ ← 2'den farklı bir şey olamaz

$\left[1/4 + 1/2\right] \cdot 1/4 = 3/4 \cdot 1/4 = \frac{3}{16}$

$2 \to HH \to 1/4$

$1 \left\{ \begin{matrix} HT \\ TH \end{matrix} \right\} 1/2$

$0 \{ TT \to 1/4$

4

---

independence of n random variables

Consider n random variables $X_1, X_2, \ldots X_n$

We say that $X_1, X_2, \ldots, X_n$ are independent

Bu ikisi eşittir

$$P(X_1 = x_1, X_2 = x_2, X_3 = x_3, \ldots, X_n = x_n)$$

$$= P(X_1 = x_1) \cdot P(X_2 = x_2) \ldots P(X_n = x_n)$$

for all $x_1, x_2 \ldots x_n$

Special distributions

olumlu veya olumsuz diyebilceğiniz iki olası sonucu olan rastgele bir deneydir

1- Bernoulli distribution

A bernoulli trial is a random experiment that has two possible outcomes which can be labeled as success of failure

Simplest discrete random variable

Başarı veya başarısızlık gibi iki olası sonucu vardır

$$P_X(x) = \begin{cases} 1 & \text{for } x = 1 \\ 0 & \text{outcomes} \end{cases}$$

indicator random variable

child — G 1

child — B 0

$$I_A = \begin{cases} 1 & \text{success} \\ 0 & \text{failure} \end{cases}$$

1 Head

$$I_A \sim \text{Bernoulli}(P(A))$$

Coin

coin — 0 tail

Definition: A random variable X is said to be a bernoulli random variable with parameter p, $X \sim Bernoulli(p)$ if its PML

$$P_X(x) = \begin{cases} p & \text{for } x=1 \\ 1-p & \text{for } x=0 \\ 0 & \text{otherwise} \end{cases}$$

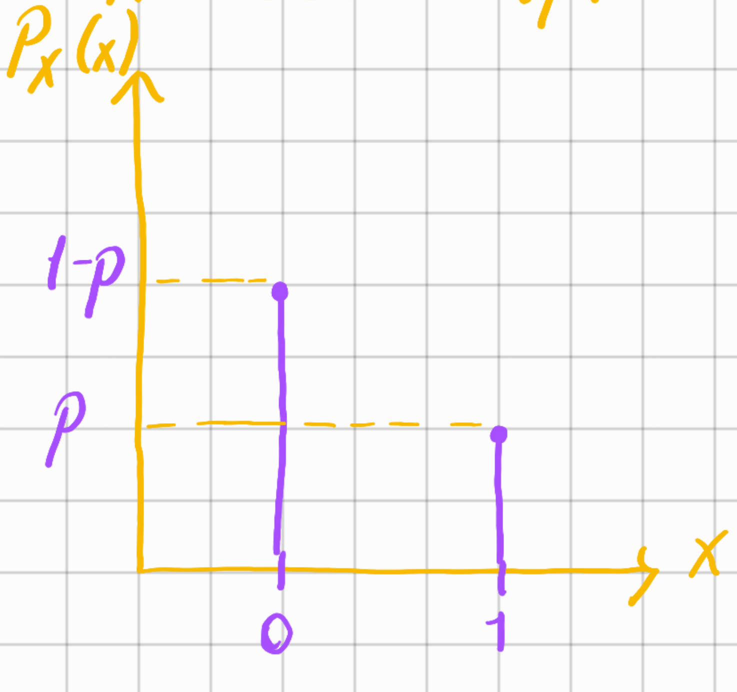
Belirli bir olayla ilişkili bu olay gerçekleşirse başarı gerçekleşmezse başarısızlık

indicator random variable;

Bernoulli random variable (BRV) is also called indicator random variable

↳ Bernoulliye indicator random variable da denir

$X \sim Bernoulli(p)$



Geometric Distribution

Suppose we have a coin

$P(H) = p$ I toss the coin until I observe the first head. We define X as the number coin tosses. X is said to have geometric distribution

Bu deneyi bağımsız bernoulli denemelerinin tekrarı olarak düşünebilirsiniz

Repeating Bernoulli trials until the first success

$\underbrace{T \ldots T}_{(k-1)\text{ tosses}} H \longrightarrow k$ tosses

$(k-1)$ tosses → Bunların her birinin olasılığı $(1-p)$

k^{th} → Bunun olasılığı p olur

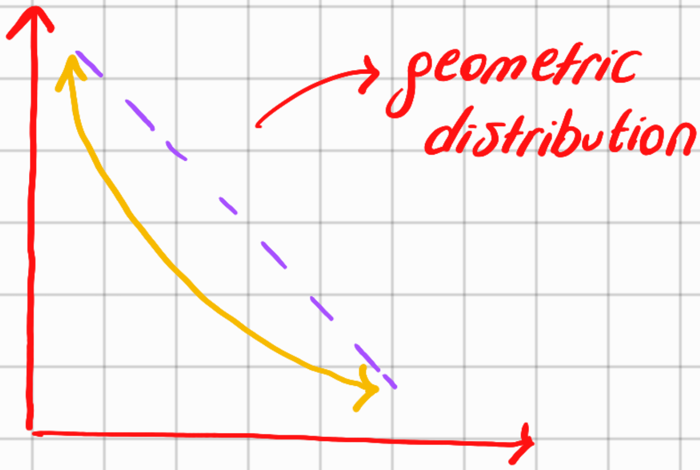
$\Big\} \; p \cdot (1-p)^{k-1}$ for $k = 1, 2, 3 \ldots$

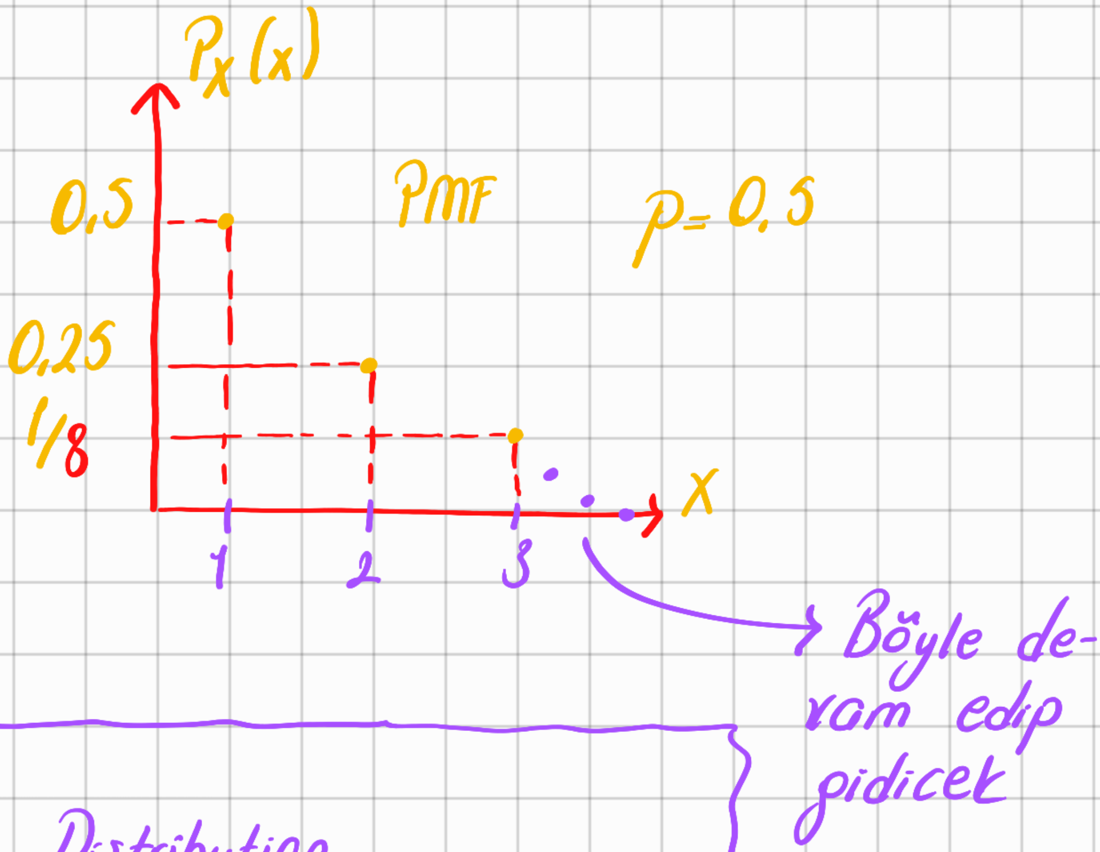
Definition

A random variable X is said to be a geometric random variable with parameter p, $X \sim Geometric(p)$ its PMF is

$P(k) = \{ \; (1-p)^{k-1} \ldots$

$$P_X(k) = \begin{cases} p.(1-p) & \text{for } k=1,2,3\ldots \\ 0 & \text{otherwise} \end{cases}$$





3_ Binomial Distribution

$P(H) = p$ toss a coin n times

geometric distributionla arasındaki fark burdaki n'yi biliyoruz

X = number of heads

$R_X = \{0, 1, \ldots, n\}$

→ n independent Bernoulli trials

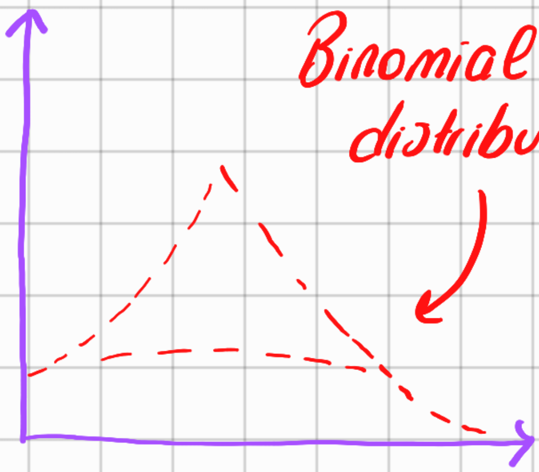
PMF of X

$$\binom{n}{k} p^k (1-p)^{n-k}$$

Definition: A random variable X is said to be binomial random variable with parameters n and p, shown as $X \sim Binomial(n,p)$ and its PMF

$$P_X(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & \text{for } k = 0, 1, \ldots, n \\ 0 & \text{otherwise} \end{cases}$$

Binomial distribution



Binomial is sum of Bernoullies

$X_1, X_2, \ldots X_n$ are independent bernoulli(p) random variables

$X = X_1 + X_2 + X_3 \ldots + X_n$ has a binomial(n,p) distribution

---

Example: $X \sim \text{Binomial}(n,p)$ and $Y \sim \text{Binomial}(m,p)$

$Z = X + Y$ find PMF of Z

$$\left.\begin{aligned} X &= X_1 + X_2 + \ldots X_n \\ Y &= Y_1 + Y_2 + \ldots Y_m \end{aligned}\right\} n+m$$

$\rightarrow$ X ve Y independent olur

$Z = X + Y = X_1 + X_2 \ldots X_n + Y_1 + \ldots Y_m$ since X_i and Y_J

are independent Bernoulli trials

$$P_Z(k) = \begin{cases} \binom{m+n}{k} p^k \underbrace{(1-p)}_{\text{tails}}^{m+n-k} & \text{for } k = 0, 1 \ldots m+n \\ 0 & \text{otherwise} \end{cases}$$

---

4. Pascal (Negative Binomial) Distribution

4- Pascal (Negative Binomial) Distribution

it is the generalization of geometric distribution (geometric distribution'ın genellemestirilmiş hali)

Burda deney başarılı olana kadar tekrarlamamız gerekiyor

Random exp. of independent trials until observing m success

Coin $P(H) = p$

toss the coin until I observe m heads

X as the number of coin tosses

X is said to have pascal distribution with parameters (p, m)

$X \sim pascal(m, p)$

if $m = 1$ $pascal(1, p) = $ Geometric

Range of X $R_X = \{m, m+1, m+2, \ldots\}$

event $A = \{X = k\}$

$A = B \cap C$ where
B event that we have $m-1$ heads in first $k-1$ trials
C event that we observe a head in k^{th} trial

$P(A) = P(B \cap C) = P(B) . P(C)$

Binomial

$P(C) = p$

$k-1-m+1$

$(k-1)-(m-1)$

$$P(B) = \binom{k-1}{m-1} p^{m-1} (1-p)^{(k-1)-(m-1)}$$

$$\binom{k-1}{m-1} p^{m-1} (1-p)^{k-m}$$

$$P(A) = P(B) \cdot P(C)$$

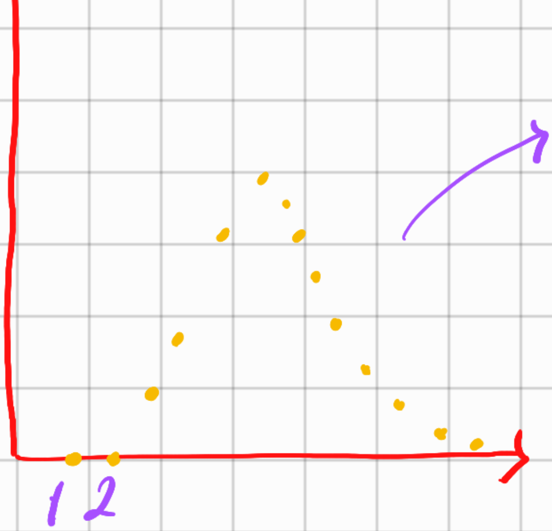
$$= \binom{k-1}{m-1} p^{m-1} (1-p)^{k-m} * p$$

$$= \binom{k-1}{m-1} p^{m} (1-p)^{k-m}$$

pascal distribution

$X \sim \text{pascal}(m, p)$

$m = 3$



plot a pascal distribution

5_ Poission Distribution

Belirli olayların oluşumlarını saydığımız durumlarda kullanılır

→ Used in cases when we are counting the occurances of certain events in an interval of time or space

→ Approximation of real-world random variable

We are counting the number of customers at a certain time period.

On average $\lambda = 15$ customers visit the store

Bunun gerçek dünyadaki rastgele değişkenin bir tahmini olduğunu söyleyebiliriz

müşteri sayısı

$\lambda > 0$

0

$n = 360$ sec

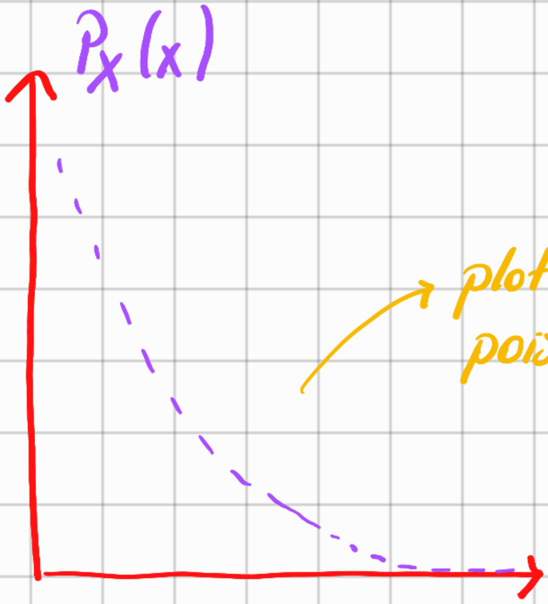
$p = \frac{\lambda}{n}$

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

$p = \frac{\lambda}{n}$ → n is very large → n çok büyükse limit ile çözebiliriz

$\lim_{n \to \infty} P(X=k) = \underbrace{\text{solve}}_{\text{Books! check}} = \underbrace{\frac{e^{-\lambda} \cdot \lambda^k}{k!}}_{\text{posision dist}}$ for $k \in R_X$

→ Bu denklemi nasıl elde ettiğini görmek için kitaba bak



→ plot of poison distribution

poison distribution perçet random variablelerin yoklaşıklığı için kullanılır

Example: Number of emails I get model poisson distribition

→ Avg. emails per minute 0.2

what is the probability that I get no emails in an interval of 5 minutes?

ⓐ X is a Poisson random variable $\lambda = 5 \cdot 0.2 \to 1$

$P(X=0) = P_X(0) = \frac{e^{-\lambda} \cdot \lambda^0}{\ldots} = \frac{e^{-1} \cdot 1}{\ldots} = \frac{1}{\ldots} = \frac{1}{\ldots}$

0! 1 e 2,71

Burdaki 0 no emailden geliyor

(b) 5 dk en az 3 email? 1-(en fazla 3'ü çıkar)

en fazla 3 email

$P(X=0)+P(X=1)+P(X=2)+P(X=3)$

Bunları formülde yerine koycaz