PRIME INTUIT
Finishing School

Bernoulli Distribution

Experiments with only two outcomes?

Bernoulli Trials

Out come: {Positive, Negative}

Outcome: {Pass, Fail}

Outcome: {Hit, Flop}

Outcome: {Spam, Ham}

Outcome: {Approved, Denied}

Ω : Failure Success

X: Ω : → {0. 1}

Bernoulli Random Variable

Bernoulli Distribution

Bernoulli Distribution?

Bernoulli Trials

Ω : Failure, Success

A: event that the outcome is success

X: Ω : → {0. 1}

Let P(A) = P(success) = P

Bernoulli Random Variable

Px(1) = P

Px(0) = 1 - P

$Px(x) = p^{x} * (1\text{-}P)^{(1-x)}$

Bernoulli distribution

Bernoulli Distribution

Is Bernoulli distribution a valid distribution ?

$\mathrm{P}x(x) = p^{x} * (1\text{-P})^{(1-x)}$ Bernoulli distribution

Ω : Failure Success

X: Ω : → {0. 1}

↑

$Bernoulli\ Random\ Variable$

$Px(x) \geq 0$

$\Sigma_{x\in\{0,1\}}Px(x) = 1$?

$\Sigma_{x\in\{0,1\}}Px(x) = Px(0) + Px(1)$

$= (1\text{-P}) + \text{P} = 1$

Binomial Distribution

Repeat a Bernoulli trial n times

….. n number of times

Independent :

(Success / failure in one trail does not affect the outcome of other trails)

I*dentical* :

(Probability of success ‘P’ in each trial is the same)

What is the probability of k successes in n trails? ($k \in [0,n]$)

Binomial Distribution

Binomial Distribution (Examples)

….. n number of times

Each ball bearing produced in a given factory is independently non defective with probability p

If you select n ball bearings what is the probability that k of them will be defective ?

What is the probability of k successes in n trails? ($k \in [0,n]$)

PRIME INTUIT
Finishing School

Binomial Distribution

Binomial Distribution (Examples)

….. n number of times

The probability that a customer purchases something from yoru website is P

Assumption 1: Customers are identical (economic strata, interests, needs, etc)

Assumption 2: Customers are independent (one’s decision does not influence another)

What is the probability of k out of n customers will purchase something ?

Binomial Distribution

Binomial Distribution (Examples)

….. n number of times

Marketing Agency: The probability that a customer opens your email is P

Assumption 1: Customers are identical

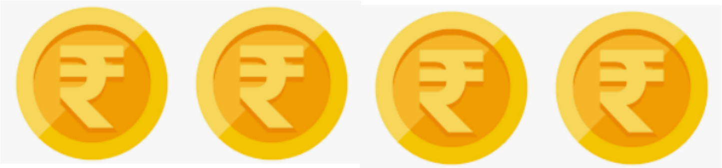
Assumption 2: Customers are independent

If you send n emails what is probability that the customer will open at least one of them ?

PRIME INTUIT
Finishing School

Binomial Distribution

Binomial Distribution



….. n number of times

How many different outcomes can we have if we repeat a Bernoulli trail n times ?

S, F

___ ___ ___ ___ ___ n

(sequence of length n from a given set of 2 objects)

= 2^noutcomes

PRIME INTUIT
Finishing School

Binomial Distribution

Binomial Distribution

….. n number of times

Ω $\qquad\qquad$ X

$T\ T\ T$
T T H
T H T
T H H
H T T
H T H
H H T
H H H

0
1
2
3

Example: n = 3, k = 1

H = Success
T = Fail

____ ____ ____

A = { HTT, THT, TTH}

$P_x(1) = P(A)$

$P_x(1) = P(A)$ = P({HTT}) + P({THT}) + P({TTH})

P({HTT}) = p (1-p) (1-p)

P({THT}) = (1-p) p (1-p)

P({TTH}) = (1-p) (1-p) p

$P_x(1) = P(A) = 3(1\text{-}p)^2\ p$

Binomial Distribution

Binomial Distribution

..... n number of times

Ω X

$T\ T\ T$
T T H
T H T
T H H
H T T
H T H
H H T
H H H

0
1
2
3

Example: n = 3, k = 1

A = { HTT, THT, TTH}

___ ___ ___

3 trials and 1 success

= 3 choose 1 = $\binom{3}{1}$

$P_x(1) = P(A)$ = 3(1-p)^2 p

= 3(1-p)$^{(3-1)}$ p^1

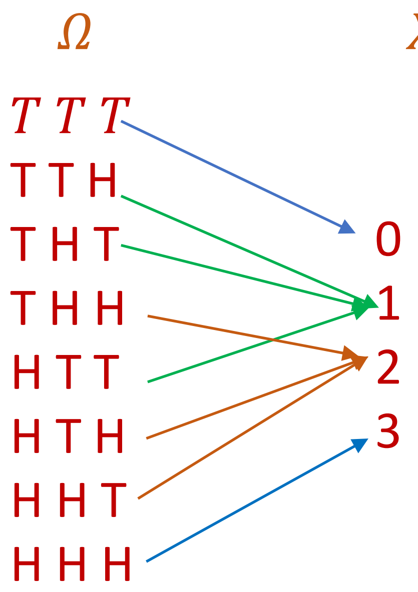
= $\binom{3}{1}$ (1-p)$^{(3-1)}$ p^1

PRIME INTUIT
Finishing School

Binomial Distribution

Binomial Distribution

….. n number of times



Example: n = 3, k = 2

B = { HTH, HHT, THH}

____ ____ ____

$P_x(2) = P(B)$ = 3(1-p) p^2

= 3(1-p)$^{(3-2)}p^2$

= $\binom{3}{2}$ (1-p)$^{(3-2)}$ p^2

3 trials and 2 success

= 3 choose 2 = $\binom{3}{2}$

Binomial Distribution

Binomial Distribution

Observations

n trials and k success

$\binom{n}{k}$ terms in the summation

$\binom{n}{k}$ favourable outcomes

each terms will have the factor p^k

$each\ of\ the\ k\ success\ occur\ independently\ with\ a\ probability\ p$

each terms will have the factor (1 – p)$^{(n-k)}$

each of the n – k failures occur independently with a probability 1 - p

$$P_x(\mathrm{k}) = \binom{n}{k}\ p^k\ (1\text{-p})^{(n-k)}$$

Parameters: p, n

The entire distribution is full specified once the value of p and n are known

PRIME INTUIT
Finishing School

Binomial Distribution

Binomial Distribution Example1: Social distancing

$Suppose\ 10\%\ of\ your\ colleagues\ from\ workplace\ are\ infected\ with\ COVID-19\ but\ are\ asymptomatic\ (hence\ come\ to\ office\ as\ usual)$

$Suppose\ you\ come\ in\ close\ proximity\ of\ 50\ of\ your\ colleagues.$
What is the probability of you getting infected

n = 50, p = 0.1

P(getting infected) = P(at least one success)

= 1 – P(0 successes)

= 1 - $P_x(0)$

= 1 - $\binom{50}{0}$ p^0 $(1\text{-}p)^{(50)}$

= $1 - 1 * 1 * 0.9^{(50)} = 0.995$

= $1 - 1 * 1 * 0.9^{(10)} = 0.6513$

= $1 - 1 * 1 * 0.98^{(10)} = 0.1829$ $,P\ change\ to\ 2\%$

Binomial Distribution

Binomial Distribution Example2: Mac Users

$Suppose\ 10\%\ of\ students\ in\ your\ class\ use\ Mac\ book,$
$If\ you\ select\ 25\ students\ at\ random$

$a)\ what\ is\ the\ probability\ that\ exactly\ 3\ of\ them\ are\ using\ Mac\ book\ ?$

b) $what\ is\ the\ probability\ that\ between\ 2\ to\ 6\ of\ them\ are\ using\ Mac\ book\ ?$

C) $How\ would\ the\ above\ probabilities\ change\ if\ instead\ of\ 10\%, 90\%\ were\ using\ Mac\ book\ ?$

a) n = 25, p = 0.1 and k = 3

b) n = 25, p = 0.1 and k ={2,3,4,5,6}

c) n = 25, p = 0.9, 0.5

PRIME INTUIT
Finishing School

Binomial Distribution

Binomial Distribution Example2: Mac Users

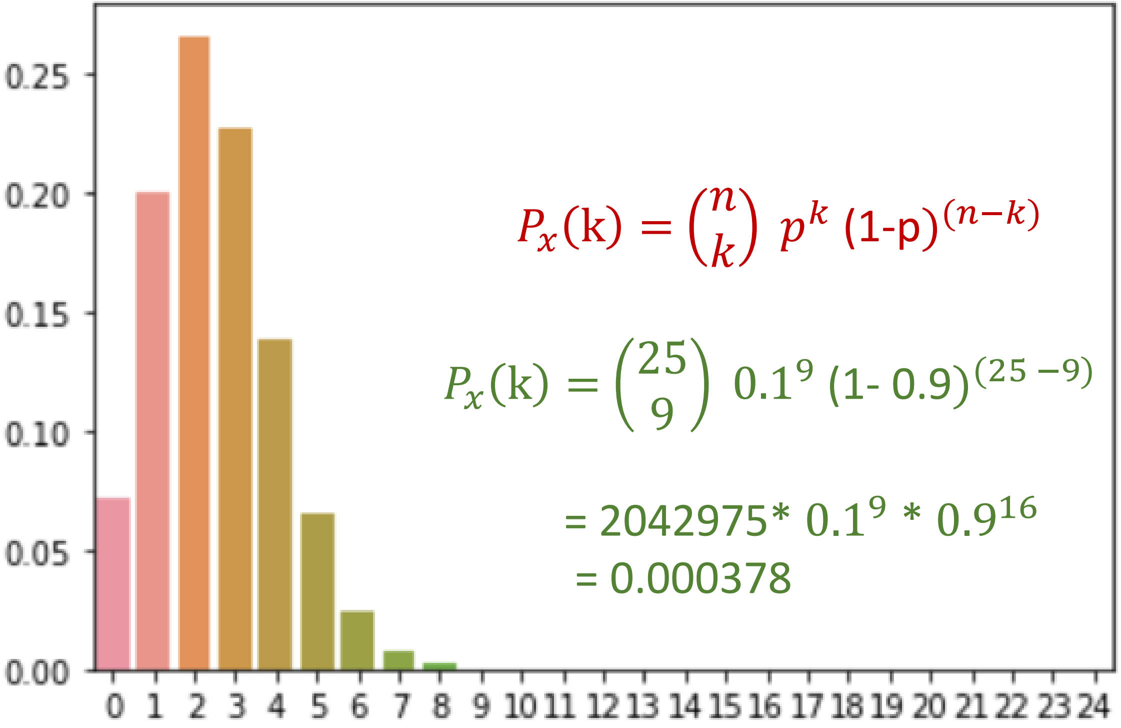
Suppose 10% of students in your class use Mac book,
If you select 25 students at random

a) n = 25, p = 0.1 and k = 3

b) n = 25, p = 0.1 and k ={2,3,4,5,6}

c) n = 25, p = 0.9, p = 0.5

```
import seaborn as sb
import numpy as np
from scipy.stats import binom

x = np.arange(0,25)
n = 25
p = 0.1

dist = binom(n, p)
ax = sb.barplot(x = x, y = dist.pmf(x))
```



$$P_x(\mathrm{k}) = \binom{n}{k}\, p^k\ (1\text{-p})^{(n-k)}$$

$$P_x(\mathrm{k}) = \binom{25}{9}\ 0.1^9\ (1\text{- }0.9)^{(25-9)}$$

$$= 2042975 * 0.1^9 * 0.9^{16}$$
$$= 0.000378$$

Binomial Distribution

Binom

Suppo
If you

a) n = 25, p = 0.1 an

b) n = 25, p = 0.1 an

c) n = 25, p = 0.9, p

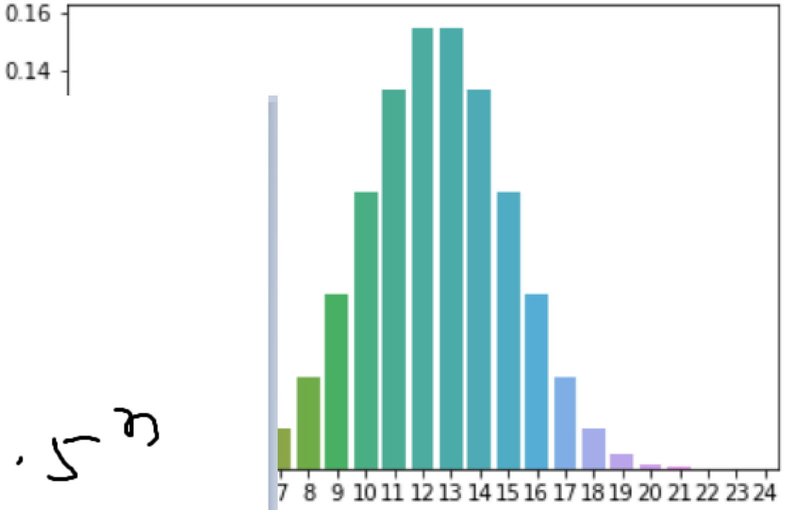
```
import seaborn as
import numpy as r
from scipy.stats

x = np.arange(0,2
n = 25
p = 0.9

dist = binom(n, p)
ax = sb.barplot(x = x, y = dist.pmf(x))
```

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

$$(0.5)^k (0.5)^{n-k} = 0.5^n$$

$$\binom{n}{k} = \binom{n}{n-k}$$

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

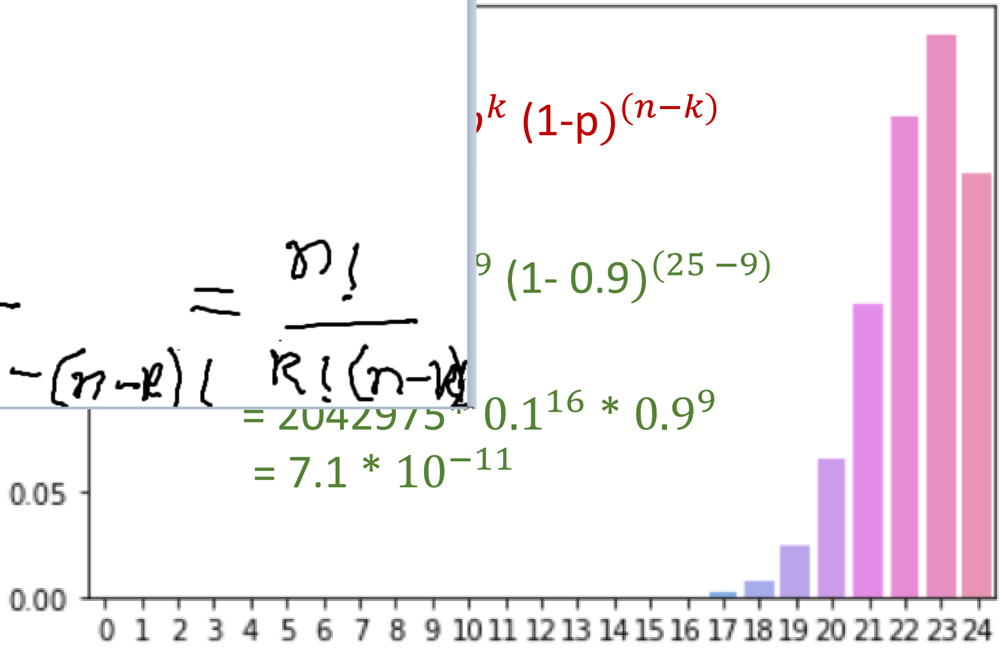


^{k} (1-p)$^{(n-k)}$

^{9} (1- 0.9)$^{(25-9)}$

= 2042975 * 0.1^{16} * 0.9^9

= 7.1 * 10^{-11}



Binomial Distribution

Is Binomial Distribution a valid distribution ?

$p_x(x) \geq 0$ $\qquad$ $P_x(\mathrm{k}) = \binom{n}{k}\, p^k\, (1\text{-p})^{(n-k)}$

$\sum_{i=0}^{n} p_x(i) = 1\ \ ?$

$= p_x(0) + p_x(1) + p_x(2) + \ldots\ldots\ldots\ldots + p_x(n)$

$= \binom{n}{0}\, p^0\, (1\text{-p})^{(n-0)} + \binom{n}{1}\, p^1\, (1-\mathrm{p})^{(n-1)} + \binom{n}{2}\, p^2\, (1-\mathrm{p})^{(n-2)} + \cdots \ldots \ldots \ldots + \binom{n}{n}\, p^n\, (1-\mathrm{p})^{(n-n)}$

$(a+b)^n = \binom{n}{\theta} a^0 b^n + \binom{n}{1} a^1 b^{n-1} + \binom{n}{2} a^2 b^{n-2} \ldots \ldots \ldots + \binom{n}{n} a^{\mathrm{n}} b^0$

$\quad on\ the\ left\ hand\ side\ if\ (a = p, \qquad b = 1 - p)\ then$

$= (\,P + 1\ - P)^n = 1^n = 1$

Binomial Distribution

Relation between Binomial and Bernoulli

Binomial $P_x(\mathrm{k}) = \binom{n}{k} p^k (1\text{-p})^{(n-k)}$

Bernoulli $n = 1, \quad k \in \{0, 1\}$

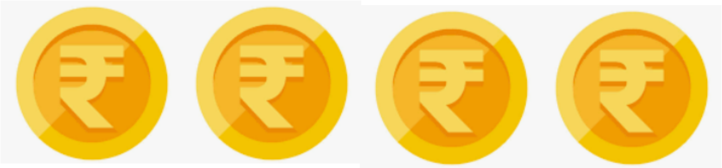
$$p_x(0) = \binom{1}{0} p^0 (1\text{-p})^{(1-0)} = 1 - \mathrm{p}$$

$$p_x(1) = \binom{1}{1} p^1 (1\text{-p})^{(1-1)} = \mathrm{p}$$

Bernoulli distribution is a special case of Binomial distribution

Geometric Distribution

Geometric Distribution



….. ∞ number of times

X : The number of tosses until we see the first heads

$R_x = \{1, 2, 3, 4, 5, \dots.\}$

___ ___ ___ ___ ___ n

Why would we be interested in such a distribution ?

(sequence of length ∞ from a given set of 2 objects)

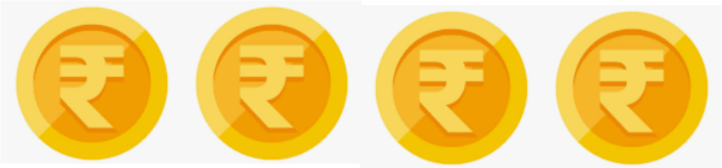
= 2^{∞}outcomes

$p_x(x)$ = ?

PRIME INTUIT
Finishing School

Geometric Distribution

Geometric Distribution



….. ∞ number of times

Why would we be interested in such a distribution ?

Because: It’s useful in any situation involving “waiting times”

Independent trials identical distribution P(success) = p

Street vendor selling vada pav outside a subway station
(Chance that the first vada pav will be sold after k trails)

Salesman handing pamphlets to passersby
(Chance that the k-th person will be the first person to actually read the pamphlet)

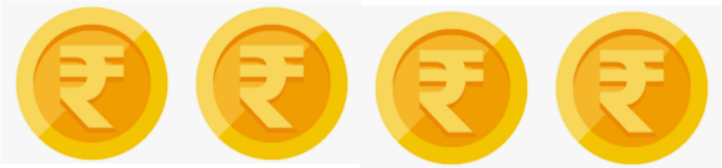
A digital marketing agency sending emails
(Chance that the k-th person will be the first person to actually read the email)

Geometric Random variable X, number of trials after which we will get the first success

PRIME INTUIT
Finishing School

Geometric Distribution

Geometric Distribution



..... ∞ number of times

P(Success) = P

$P_x(\mathrm{x})$

Example: k = 5

$P_x(5)$

F F F F S

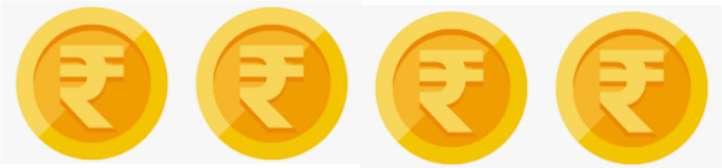
$$\underbrace{(1-p)\,(1-p)\,(1-p)\,\,(1-p)\,\,(1-p)}_{(5\ -1)}\ \underbrace{p}_{1}$$

$P_x(5)$ = (1- p)$^{(5-1)}$ p

$P_x(\mathrm{k})$ = (1- p)$^{(k-1)}$ p

Geometric Distribution

Geometric Distribution



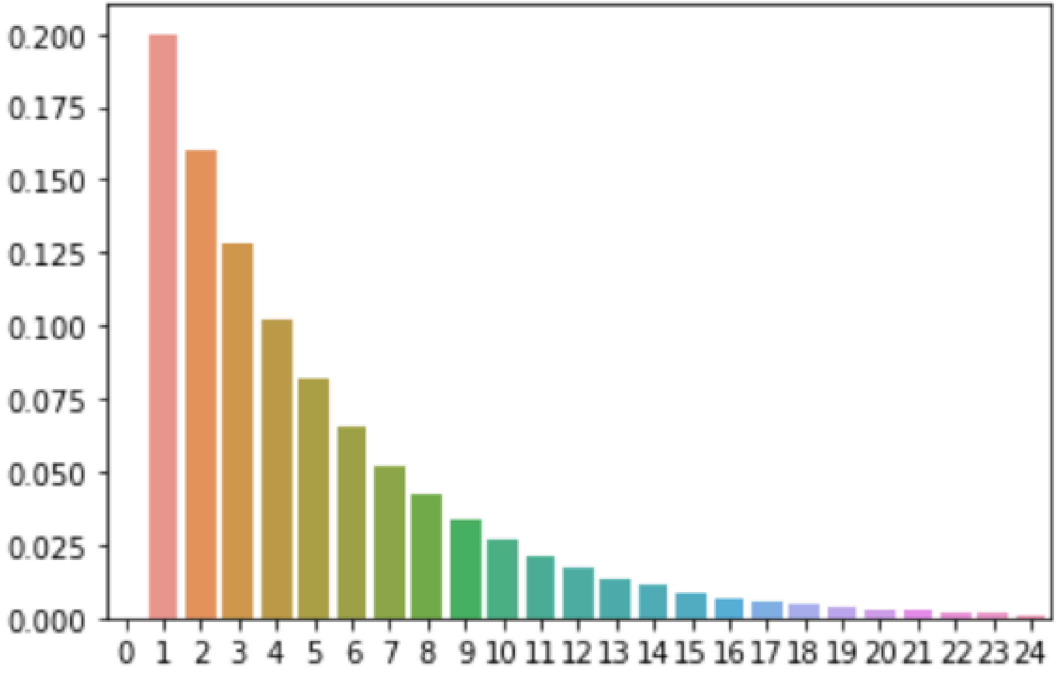
….. ∞ number of times

P (success) = p

```
import seaborn as sb
import numpy as np
from scipy.stats import geom

x = np.arange(0,25)
n = 25
p = 0.2

dist = geom(p)
ax = sb.barplot(x = x, y = dist.pmf(x))
```

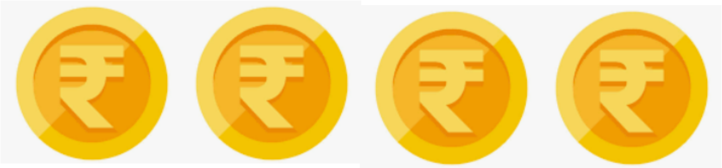
Example: p = 0.2, n = 25

$$P_x(k) = (1-p)^{(k-1)}\, p$$



Geometric Distribution

Geometric Distribution



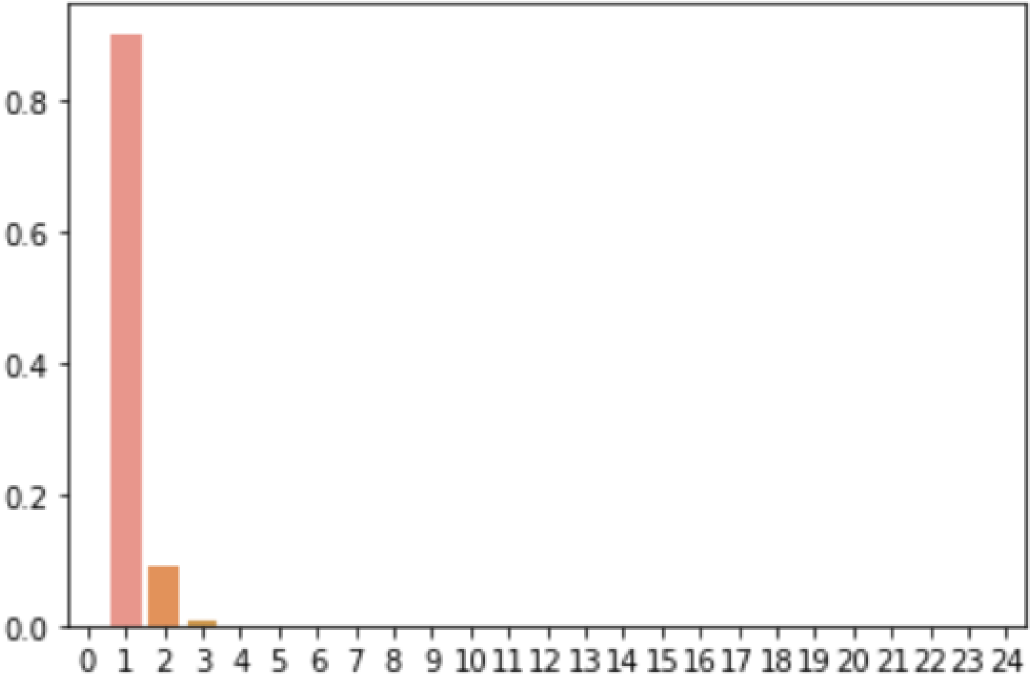
….. ∞ number of times

P (success) = p

```
import seaborn as sb
import numpy as np
from scipy.stats import geom

x = np.arange(0,25)
n = 25
p = 0.9

dist = geom(p)
ax = sb.barplot(x = x, y = dist.pmf(x))
```

Example: p = 0.9, n = 25

$$P_x(k) = (1 - p)^{(k-1)}\, p$$



Geometric Distribution

Geometric Distribution

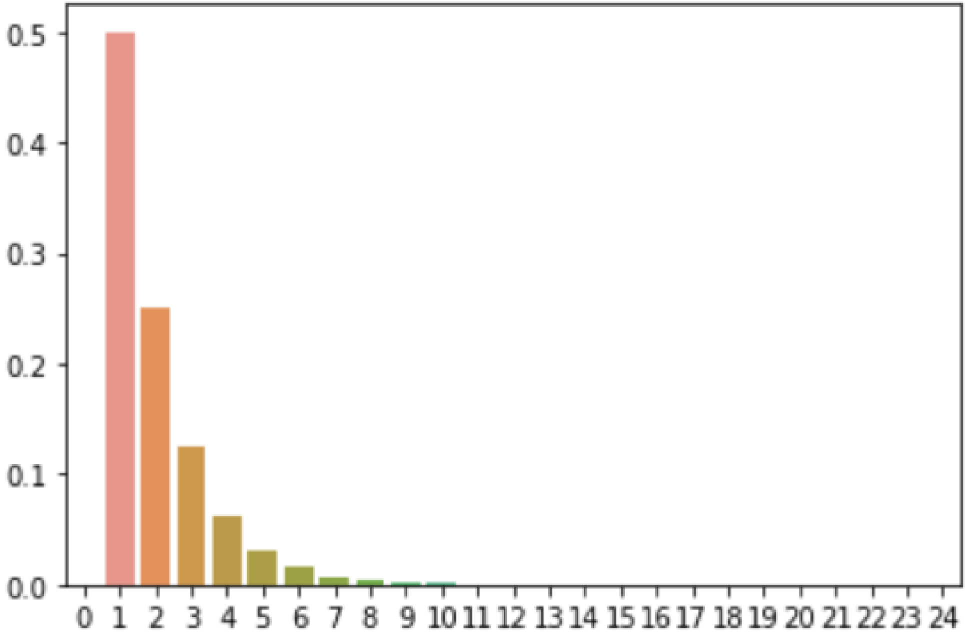
….. ∞ number of times

P (success) = p

```python
import seaborn as sb
import numpy as np
from scipy.stats import geom

x = np.arange(0,25)
n = 25
p = 0.5

dist = geom(p)
ax = sb.barplot(x = x, y = dist.pmf(x))
```

Example: p = 0.5, n = 25

$P_x(\mathrm{k}) = (1-\mathrm{p})^{(k-1)}\ \mathrm{p}$



Geometric Distribution

Is Geometric Distribution a valid distribution ?

$p_x(x) \geq 0$ $P_x(\text{k}) = (1\text{- p})^{(k-1)}\ \text{p}$

$\sum_{\text{k}=1}^{\infty} p_x(k) = 1\ \ ?$

$= p_x(0) + p_x(1) + p_x(2) + \ldots\ldots\ldots\ldots + p_x(\infty)$

$= (1\text{-p})^0\ p + (1-\text{p})^1\ p + (1-\text{p})^2\ p + \cdots \ldots \ldots \ldots + (1-\text{p})^{\infty}\ p$

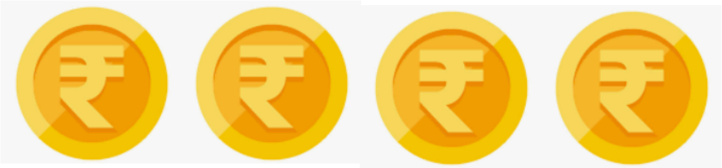
$\sum_{\text{k}=1}^{\infty} (1-\text{p})^k\ p$

$a, ar, ar^2, ar^3, ar^4, ar^5, \ldots \ldots \ldots \ldots \ldots \ldots ar^{\infty}$ $(Geometric\ series)$

$= \frac{a}{1-r} = \frac{p}{1-(1-p)} = 1$

Geometric Distribution

Geometric Distribution



….. ∞ number of times

P (success) = p

A Patient needs a certain blood group which only 9 % of the population has ?

What is the probability that the 7th volunteer that the doctor contacts will be the first one to have a matching blood group ?

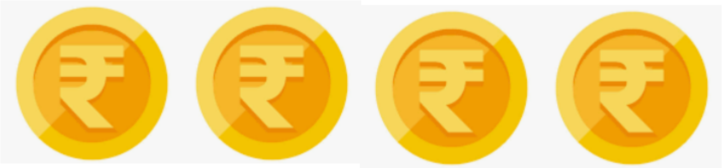
What is the probability that atleast one of the first 10 volunteers will have a matching blood type ?

$$p_x(7) = ?$$

P(x < = 10) = ?

Geometric Distribution

Geometric Distribution



..... ∞ number of times

P (success) = p

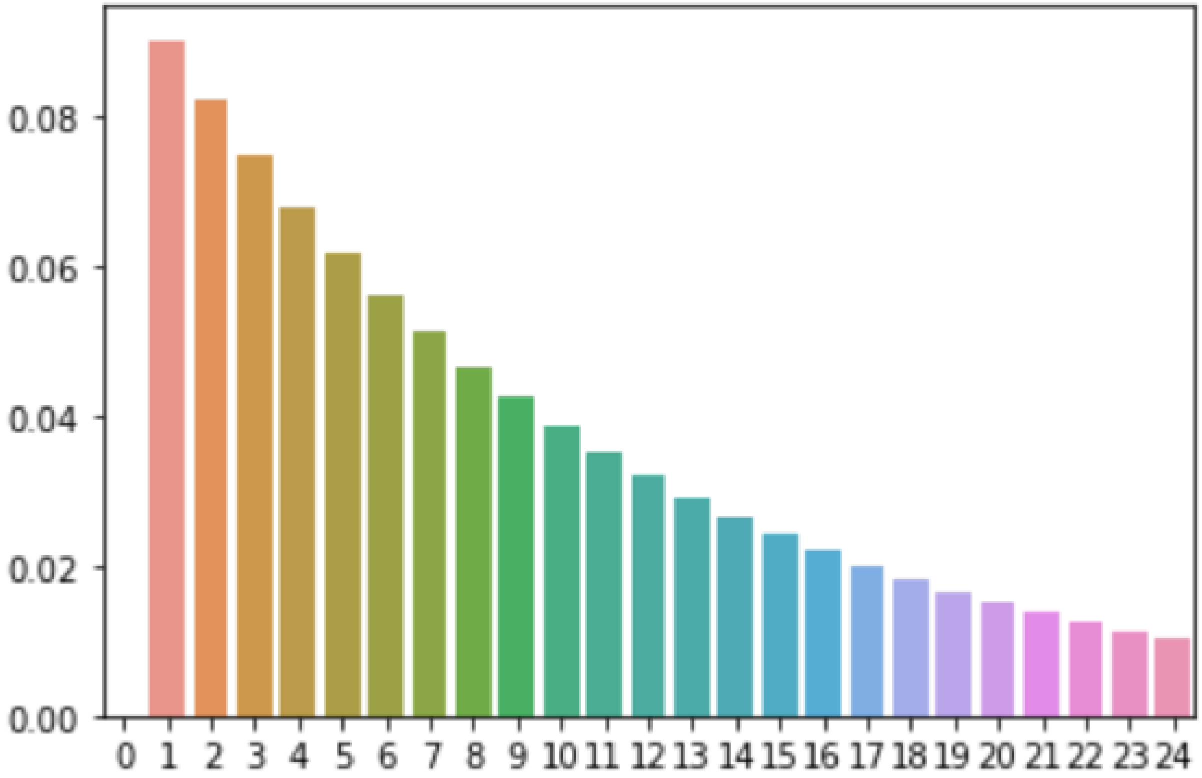
```
import seaborn as sb
import numpy as np
from scipy.stats import geom

x = np.arange(0,25)
n = 25
p = 0.09

dist = geom(p)
ax = sb.barplot(x = x, y = dist.pmf(x))
```

1) $p_x(7) = ?$

Example: p = 0.09, n = 25

2) $P(x <= 10) = 1 - P(x > 10)$

$= 1 - (1 - p)^{10}$



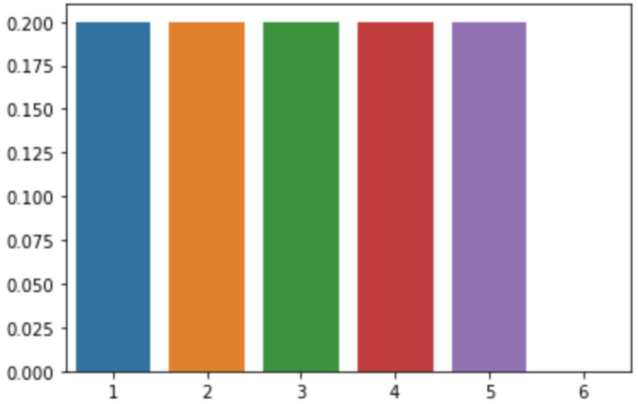
Uniform Distribution

Uniform Distribution

Experiments with equally likely outcomes

X : outcome of a die

$p_x(x)$= 1/6 $\quad \forall x \in \{1, 2, 3, 4, 5, 6\}$



X : outcome of a bingo / housie draw

$p_x(x)$= 1/100 $\quad 1 \leq \text{x} \geq 100\}$

$R_x = \{\, a \leq \text{x} \geq \text{b}\}$

Uniform Distribution

Uniform Distribution

Experiments with equally likely outcomes

X : outcome of a die / bingo / housie draw

$$p_x(x) = \frac{1}{b-a+1} \qquad a \leq x \leq b$$

Special Cases

$a = 1, b = n$

$$p_x(x) = \frac{1}{b-a+1} = \frac{1}{n} \qquad 1 \leq x \leq n$$

$a = c, b = c$

$$p_x(x) = \frac{1}{b-a+1} = \frac{1}{c-c+1} = 1 \qquad x = c$$

Uniform Distribution

Is Uniform Distribution a valid distribution ?

$$p_x(x) \leq 0, \quad p_x(x) = \frac{1}{b-a+1} \qquad a \leq x \leq b$$

$$\sum_{k=1}^{\infty} P_x(i) = 1\ ?$$

$$\sum_a^b P_x(i) = (b-a+1)\frac{1}{b-a+1} = 1$$

Expectation

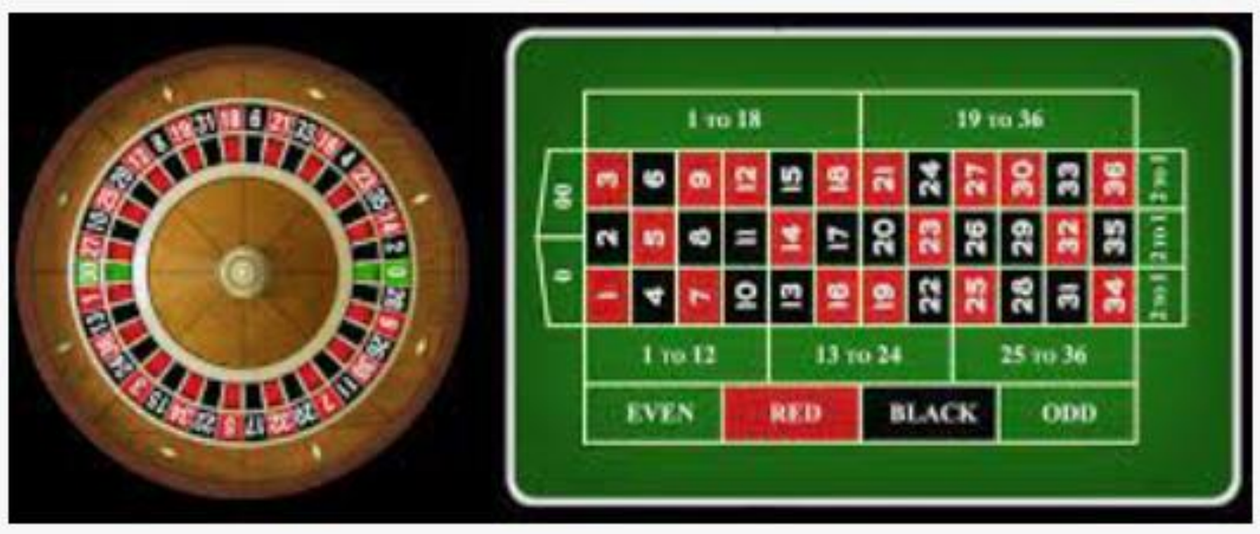
What is Expectation ?

Mathematical **expectation**, also known as the expected value, is the summation or integration of a possible values from a random variable. It is also known as the product of the **probability** of an event occurring, denoted P(x), and the value corresponding with the actual observed occurrence of the event

Expectation

Does gambling pay off ?

Roulette



$0, 0\ , 0, 1, 2, 3\ \dots\dots\dots 36$

$Standard\ pay\ off$
$= \mathbf{35{:}\,1}, if\ the\ ball\ lands$
$on\ your\ number$

$If\ you\ play\ this\ game\ a$
$1000\ times, howmuch\ do$
$you\ expect\ to\ win\ on\ an$
$average$

$if\ \ x\ is\ my\ profit\ random\ variable\ then$ (-1, 34)

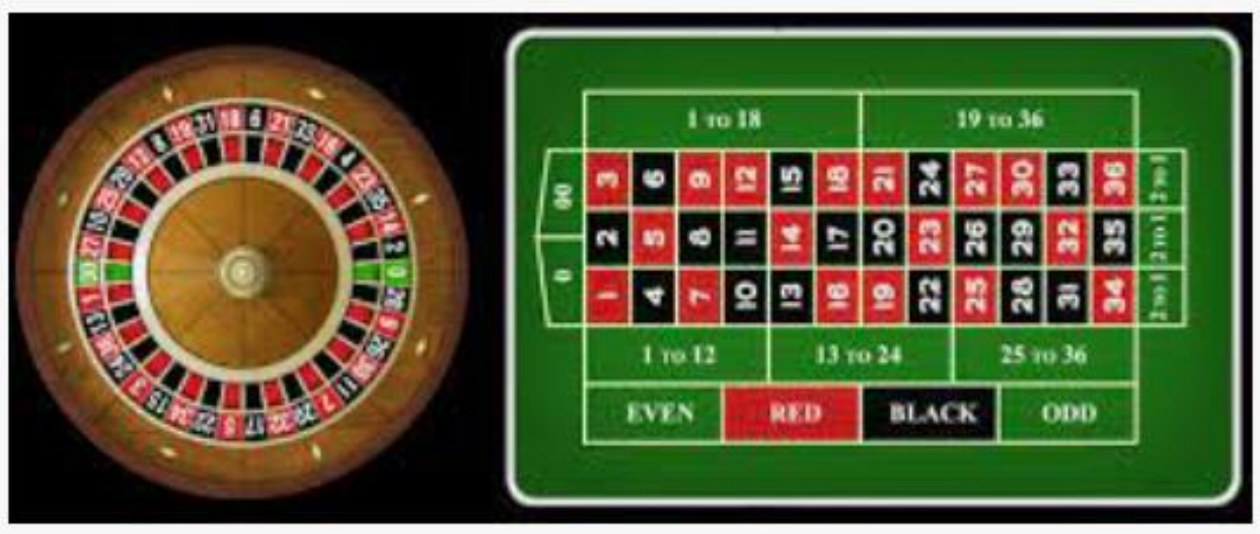
$\Omega\ is\ 00, 0, 1, 2, 3, 4, 5, 6, \dots\dots\dots\dots\dots\dots\dots\dots .36$

PRIME INTUIT
Finishing School

Expectation

Does gambling pay off ?

Roulette



if x is my profit random variable then (-1, 34)

Ω *is* $00, 0, 1, 2, 3, 4, 5, 6, \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots .36$

$Avg\ Gain = \frac{1}{1000}(26 * 34) + (974 * (-1))$ = - 90

$0, 0, 0, 1, 2, 3 \ldots \ldots \ldots 36$

$p_x(34)$ = 1/38 = 0.026

$p_x(1)$ = 1- 0.026 = 0.974

If you play this game a 1000 times, howmuch do you expect to win on an average

$p_x(win) = \frac{\#\ wins}{\#\ Games}$

$0.026 = \frac{\#\ wins}{1000}$

$\#\ wins = 26$

Expectation

Formula for Expectation:

$$E[X] = \Sigma_{x \in R_x} x * p_x(x)$$

$$E[X] = \frac{1}{1000}(\mathbf{26} * \mathbf{34}) + (\mathbf{974} * (-\mathbf{1}))$$

$$E[X] = \frac{26}{1000} * \mathbf{34} + \frac{974}{1000} * -1$$

$$E[X] = 0.026 * \mathbf{34} + 0.974 * -1$$

$$E[X] = p_x(34) * \mathbf{34} + p_x(-1) * -1$$

$$E[X] = \Sigma_{x \in \{-1,34\}} x * p_x(x)$$

$\frac{1}{2}$ $\frac{1}{2}$

$\frac{34}{2}$ $\frac{-1}{2}$

$\frac{1}{2}$ 34 + $\frac{1}{2}$-1

$p_x(34) * 34 + p_x(-1)$-1

The expected value or expectation of a discrete Random variable X whose possible values are X1, x2, x3, x4, x5, ...x_n is denoted by E[X] and computed as $E[X] = \Sigma_{i=1}^{n} x_i \, P(X = x_i)$

$$= \boldsymbol{\Sigma_{i=1}^{n} x_i * p_x(x_i)}$$

PRIME INTUIT
Finishing School

Expectation

Expectation: Insurance

A person buys a car theft insurance policy of INR 200000 at an annual premium of INR 6000. There is a 2% chance that the car may get stolen

What is the expected gain of the insurance company at the end of 1 year ?

$E[X] = \Sigma_{x \in R_x} x * p_x(x)$

$E[X] = 0.98 * 6000 + 0.02 * (-194000)$

$E[X] = 2000$

X : Profit
X : {6000, - 194000}

$p_x(6000) = 0.98$
$p_x(-194000) = 0.02$

Expectation

Expectation: Insurance

A person buys a car theft insurance policy of INR 200000, Suppose there is a 10% chance that the car may get stolen

What should the premium be so that the expected gain is INR 2000 ?

$E[X] = \Sigma_{x \in R_x} x * p_x(x)$

$E[X] = 0.9 * x + 0.1 * (x - 200000)$

$E[X] = 2000 = 0.9 * x + 0.1 * (x - 200000)$

$X = 22000$

X : Profit

X : {x, -(200000 – x)}

X : {x, (x - 200000)}

$p_x(x) = 0.90$
$p_x(x - 200000) = 0.10$

Expectation

Properties of Expectation:

Linearity of expectation $\qquad$ Y = aX + b

$E[Y] = \Sigma_{x \in R_x} g(x) * p_x(x)$

$E[Y] = \Sigma_{x \in R_x} (ax + b) * p_x(x)$

$E[Y] = \Sigma_{x \in R_x} a * x * p_x(x) + \Sigma_{x \in R_x} b * p_x(x)$

$E[Y] = a * \Sigma_{x \in R_x} x * p_x(x) + b * \Sigma_{x \in R_x} p_x(x)$

$E[Y] = a * E[X] + b * 1 \quad as \ \left(\Sigma_{x \in R_x} p_x(x)\right) = 1$

$E[X] = a\ E[X] + b$

Expectation

Properties of Expectation:

Given a set of random variables X1, X2, X3, X4…….., Xn

$$E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i]$$

We will study more about it when we do central limit theorem

PRIME INTUIT
Finishing School

Expectation

Properties of Expectation: Expectation as mean of population

Ω

W: Weights 30 45 60

N Students

$p_w(w_i) = \frac{1}{n}$

E[W] = $\Sigma_{i=1}^{n} p_w(w_i) * w_i$

E[W] = $\frac{1}{n} * w_1 + \frac{1}{n} * w_2 \dots\dots\dots \frac{1}{n} * w_n$

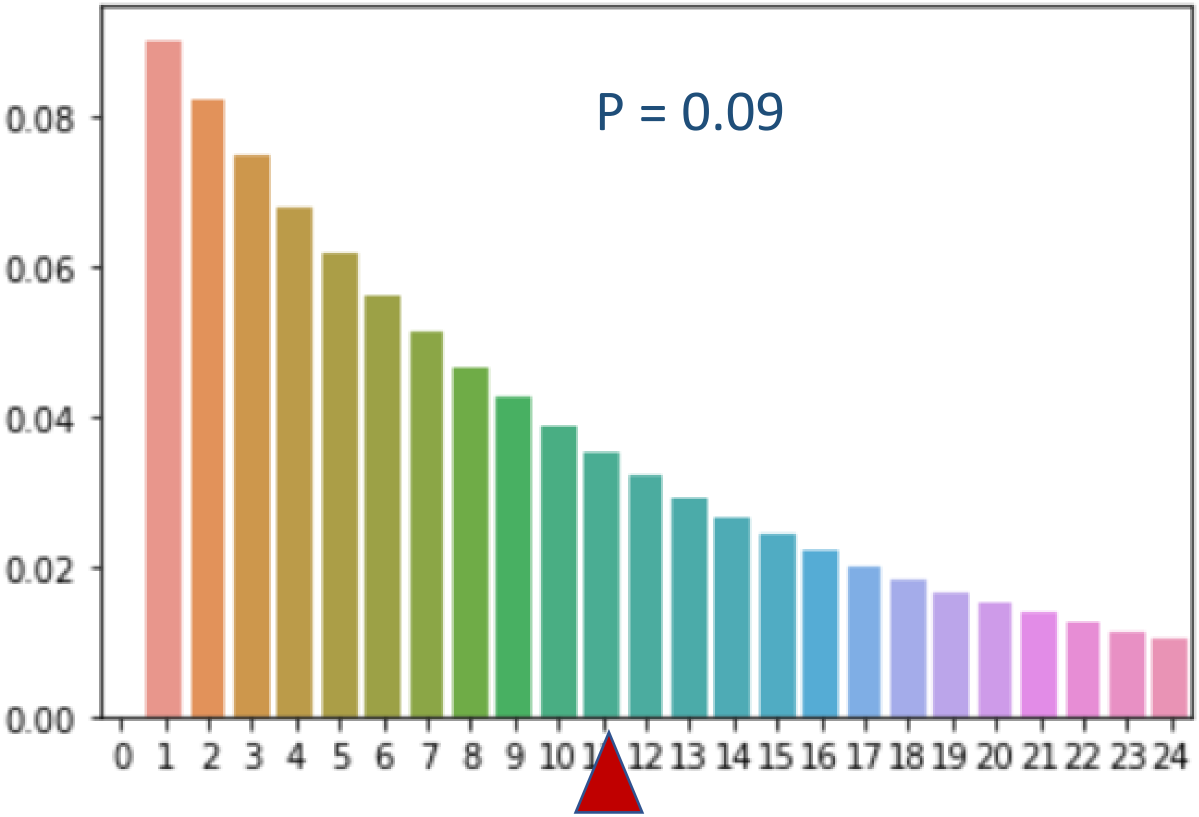
E[W] = $\frac{1}{n}\ \Sigma_{i=1}^{n} w_i$

(Center of gravity)

Expectation

Properties of Expectation: Expectation as mean of population

A patient needs a certain blood group with only 9% of the population has ?



E[X] = ?

$= p + (1-p)^1\, p + (1-p)^2\, p + \cdots \ldots \ldots \ldots + (1-p)^\infty\, p$

$= 1 * 0.09 + 2 * 0.91 * 0.09 + 3 * 0.91^2 * 0.09 + 4 * 0.91^3 * 0.09 + \ldots\ldots\ldots\ldots$

$= 0.09\,(1 + 2 * 0.91 + 3 * 0.91^2 + 4 * 0.91^3 + \ldots\ldots\ldots\ldots$

$= 0.09 \left(\frac{a}{1-r} + \frac{dr}{(1-r)^2}\right) \quad (a = 1, d = 1, r = 0.91)$

$= \frac{1}{P} = \frac{1}{0.09} = 11.11$