PRIME INTUIT
Finishing School

Discrete distributions

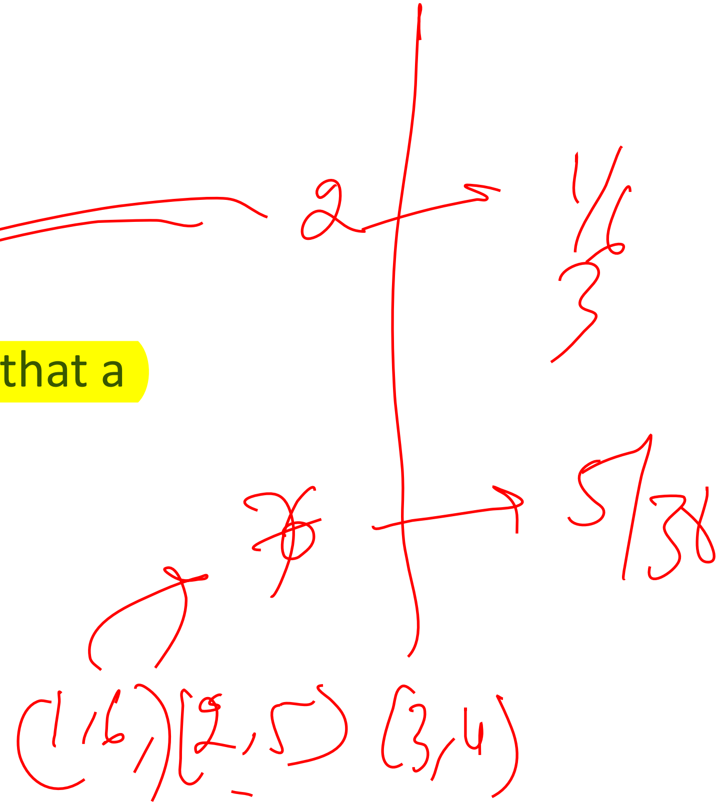
Probability Mass Functions for discrete random variables

Recap

X :	Ω	→	R
Function	**Domain**		**Range**
(Random Variable)			(Could be a subset of R)

Distribution of a random variable

An assignment of probabilities to all possible values that a discrete RV can take



(can be tedious even in simple cases)

PRIME INTUIT
Finishing School

Discrete distributions

Can PMF be specified compactly

Recap

X	P(X = x)
2	1/36
3	2/36
4	3/36
5	4/36
6	5/36
7	6/36
8	5/36
9	4/36
10	3/36
11	2/36
12	1/36

$Px(x)$ =

P(X = x)	X
1/36	If x = 2
2/36	If x = 3
3/36	If x = 4
4/36	If x = 5
5/36	If x = 6
6/36	If x = 7
5/36	If x = 8
4/36	If x = 9
3/36	If x = 10
2/36	If x = 11
1/36	If x = 12

(can be tedious when x can take on a large value)

PRIME INTUIT
Finishing School

Discrete distributions

Can PMF be specified compactly ?

P: Probability of heads
X: Random variable indicating the number of tosses after which you observe the first heads
An assignment of probabilities to all possible values that a discrete RV can take

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

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

0.5

Compact
Easy to compute
No enumeration needed

$Px(x)$ =

X
If x = 1
If x = 2
If x = 3
If x = 4
If x = 5
If x = 6
If x = 7
If x = 8
If x = 9
………
……….
If x = ∞

$Px(x)$ = (1-P)$^{(x-1)'}$ * P

1

2

10^x

11^{th}

$x=3$

$Px(3) = (1-0.5)^{(3-1)} * 0.5$

PRIME INTUIT
Finishing School

Discrete distributions

Can PMF be specified compactly ?

P: Probability of heads
X: Random variable indicating the number of tosses after which you observe the first heads

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

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

How did we arrive at the above formula ?

Is it a valid PMF (satisfying properties of a PMF)

What is the intuition behind it ?

(We will return back to this questions latter)

For now, the important point is :

It is desirable to have the entire distribution be specified by one or few parameters

Discrete distributions

Why is this important?

The entire distribution can be specified by some parameters

Cat ? Dog ? Owl ? Lion ?

P (label = cat | image) ?

$Px(x)$ = f (x)

A very complex function whose parameters are learnt from data

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

$X: \Omega : \rightarrow \{0, 1\}$

Bernoulli Random Variable

Bernoulli Distribution

Bernoulli Distribution?

Bernoulli Trials

Ω : Failure, Success

0 1

X: Ω : → {0. 1}

Bernoulli Random Variable

A: event that the outcome is success

Let P(A) = P(success) = P 0.5

$P_x(1) = P$ = 0.5

$P_x(0) = 1 - P$ = 1 - 0.5 = 0.5

$\boxed{P_x(x) = p^{x} * (1-P)^{(1-x)}}$

Bernoulli distribution

x =

strong
medi
weak

x = 3

p = 1/3

Bernoulli Distribution

Is Bernoulli distribution a valid distribution ?

$P_x(x) = p^x * (1-P)^{(1-x)}$ Bernoulli distribution

Ω : Failure Success

$X: \Omega :\rightarrow \{0. 1\}$

↑

Bernoulli Random Variable

$P_x(x) \geq 0$

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

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

$= (1-P) + P = 1$

PRIME INTUIT
Finishing School

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)

Identical :

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

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

$k \in [0,10]$

Binomial Distribution

Binomial Distribution (Examples)

….. n number of times

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

p $n=10$

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

$k=4$

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

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

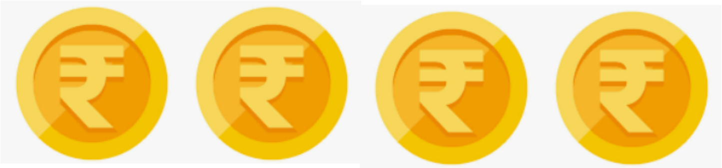
Assumption 2: Customers are independent

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

fan

Binomial Distribution

Binomial Distribution



HHTT
HHT

….. n number of times

$1 \rightarrow 2$
$2 \rightarrow 4$
$3 \rightarrow 8$
$n \rightarrow 2^n$

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

Binomial Distribution

Binomial Distribution

..... n number of times

$k=1$

Ω | X
TTT → 0
TTH → 1
THT → 1
THH → 2
HTT → 1
HTH → 2
HHT → 2
HHH → 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})

$p = 1/6$

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-p)^2 p

Binomial Distribution

Binomial Distribution

..... n number of times

Ω $\quad\quad$ X

$TTT \to 0$
TTH $\to$ 1
THT $\to$ 1
THH $\to$ 2
HTT $\to$ 1
HTH $\to$ 2
HHT $\to$ 2
HHH $\to$ 3

Example: n = 3, k = 1

A = { HTT, THT, TTH}

___ ___ ___

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

3 trials and 1 success

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

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

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

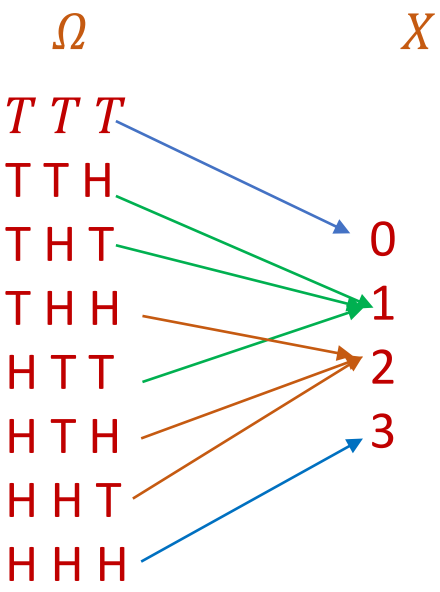
$\binom{n}{k} = {}^nC_k$

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

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

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

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

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)

p=0.

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) $\boxed{= \text{P(at least one success)}}$

= 1 – P(0 successes)

= 1 - $P_x(0)$

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

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

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

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

Binomial Distribution

Binomial Distribution Example2: Mac Users

$\frac{10}{100} = 0.1$

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

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

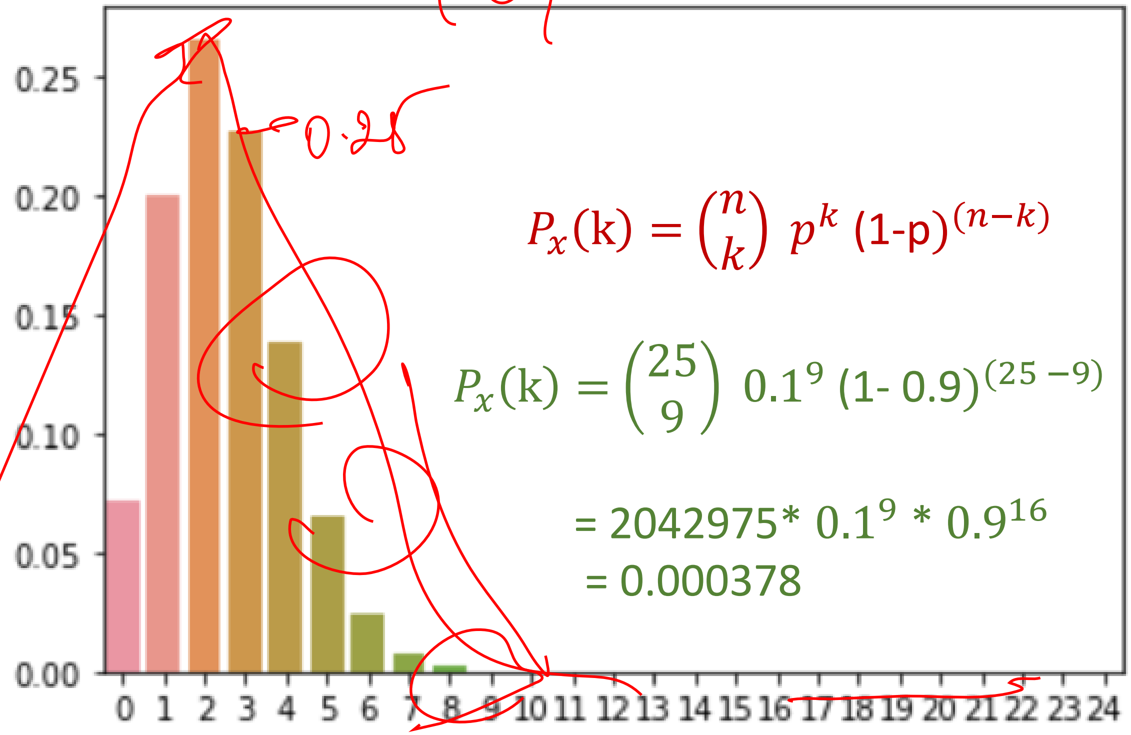
b) n = 25, p = 0.1 and k ={2,3,4,5,6}

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

```python
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))
```



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

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

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

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))
```

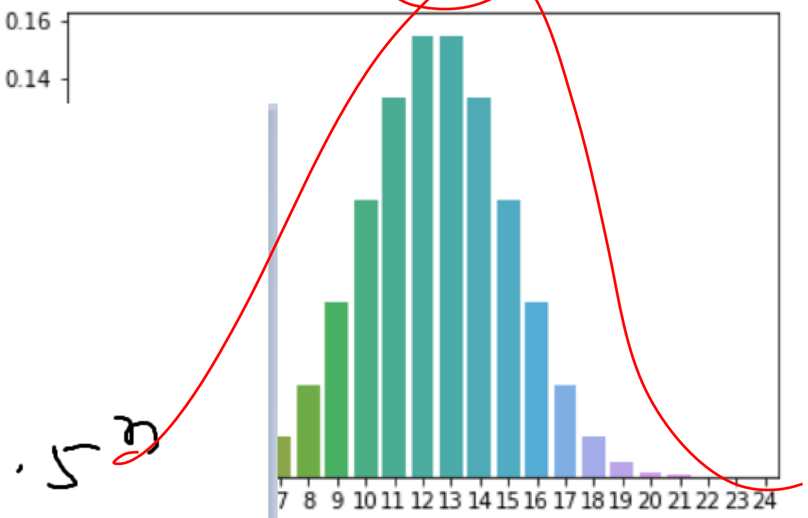
$$P_x(k) = \binom{n}{k} p^k (1-p)^{(n-k)}$$

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

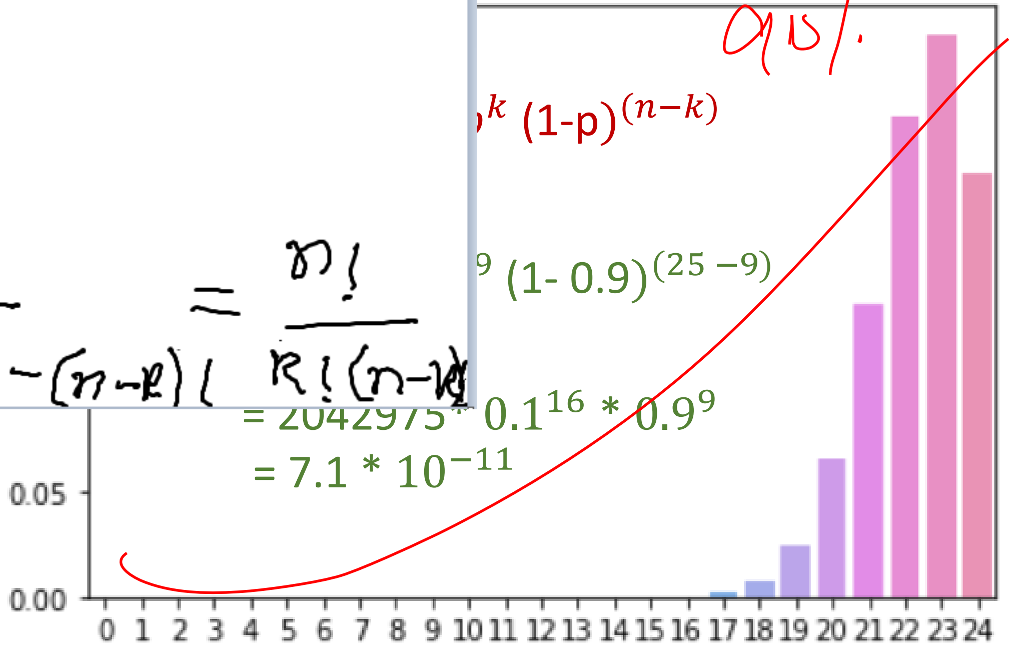
$$\binom{n}{k} = \binom{n}{n-k}$$

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

50%



95%



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

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

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

= 7.1 * 10^{-11}

PRIME INTUIT
Finishing School

Binomial Distribution

Is Binomial Distribution a valid distribution ?

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

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

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

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

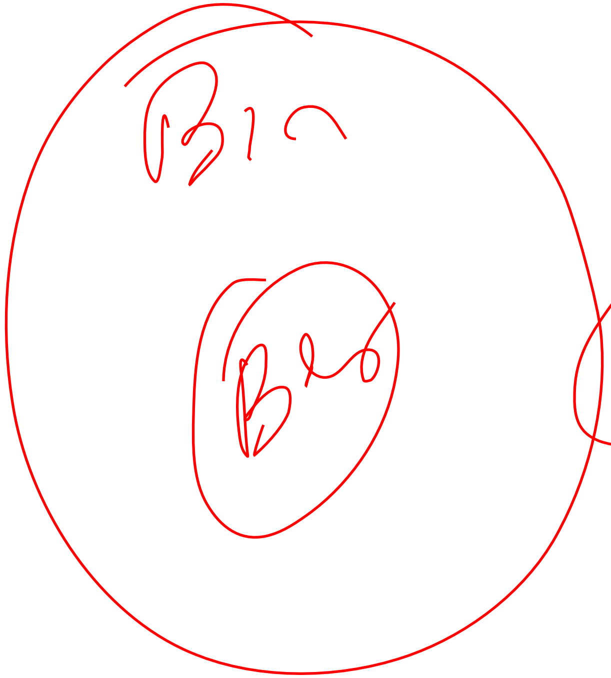
$$(a+b)^n = \binom{n}{\theta} a^0 b^n + \binom{n}{1} a^1 b^{n-1} + \binom{n}{2} a^2 b^{n-2} \ldots\ldots\ldots + \binom{n}{n} a^n b^0$$

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

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

Binomial Distribution

Relation between Binomial and Bernoulli



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

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

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

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

Bernoulli distribution is a special case of Binomial distribution