Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:07

# Probability distributions

[Next 3 classes]

- simple & popular
- what
- why
- PMF ; PDF ; CDF
  PPF ...
- real-world example(s)
- intuition
- where?

1/3

Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:06

Ops:

1. Topic → chat
   Flow
   @ end of topic

2. off-Topic / Audio → end @ Session

3. Some topics next 2-classes

2 / 3

Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:30

real-world

DS

Dalā → p/D, p/ID, arrived

Max-profit

Airlines → overbook

5-7%

Cap: 100

110 ?

risk: 110

102

2 passengers

10k $

penally

Some passengers will miss

Model using simple prob

Task: recommend the optimal # bookings

3 / 11

Chrome File Edit View History Bookmarks Profiles Tab Window Help

ProbabilityDisb_1.ipynb - Colab

colab.research.google.com/drive/1JZ-TbmRIY8mmD018C_grjyQHluBGNQxZ#scrollTo=zeaTTNEFWzxG

+ Code + Text

RAM
Disk

Linux Servers

[7]
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

import seaborn as sns
from scipy import stats
```

[8]
```
id = "1PazlhissU63pozkOjJckyjIJuM_u-0JF"
print("https://drive.google.com/uc?export=download&id=" + id)
```

```
https://drive.google.com/uc?export=download&id=1PazlhissU63pozkOjJckyjIJuM_u-0JF
```

```
!wget "https://drive.google.com/uc?export=download&id=1PazlhissU63pozkOjJckyjIJuM_u-0JF" -O flight.csv
```

Terminal /Shell /Unix

0s completed at 21:15

Flight.csv

Show all

Chrome File Edit View History Bookmarks Profiles Tab Window Help

ProbabilityDisb_1.ipynb - Colab

colab.research.google.com/drive/1JZ-TbmRIY8mmD018C_grjyQHluBGNQxZ#scrollTo=SxnxYGBUXIQp

+ Code + Text

RAM Disk

[9]

```
Resolving doc-14-14-docs.googleusercontent.com (doc-14-14-docs.googleusercontent.com)... 142.250.141.132, 2607:f8b0:
Connecting to doc-14-14-docs.googleusercontent.com (doc-14-14-docs.googleusercontent.com)|142.250.141.132|:443... co
HTTP request sent, awaiting response... 200 OK
Length: 12175 (12K) [text/csv]
Saving to: 'flight.csv'

flight.csv          100%[===================>]  11.89K  --.-KB/s    in 0s

2022-03-29 15:47:27 (55.5 MB/s) - 'flight.csv' saved [12175/12175]
```

```
flights = pd.read_csv('./flight.csv')
flights.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1000 entries, 0 to 999
Data columns (total 3 columns):
 #   Column        Non-Null Count  Dtype
---  ------        --------------  -----
 0   Passenger_ID  1000 non-null   int64
 1   Flight_ID     1000 non-null   object
 2   Arrived       1000 non-null   int64
dtypes: int64(2), object(1)
memory usage: 23.6+ KB
```

P-ID F-ID Arrived

%

0s completed at 21:19

Flight.csv

5 / 5

Show all

Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:21

Dalā : —

6 / 6

Task: how much overbooking can we do to max profit.

Chrome File Edit View History Bookmarks Profiles Tab Window Help
ProbabilityDisb_1.ipynb - Colab
colab.research.google.com/drive/1JZ-TbmRIY8mmD018C_grjyQHluBGNQxZ#scrollTo=dxAtc_fhXvov
Update
+ Code + Text
RAM
Disk

[10]

```
RangeIndex: 1000 entries, 0 to 999
Data columns (total 3 columns):
 #   Column        Non-Null Count  Dtype
---  ------        --------------  -----
 0   Passenger_ID  1000 non-null   int64
 1   Flight_ID     1000 non-null   object
 2   Arrived       1000 non-null   int64
dtypes: int64(2), object(1)
memory usage: 23.6+ KB
```

p-ID, f-ID →

```
flights.head()
```

| | Passenger_ID | Flight_ID | Arrived |
|---|---|---|---|
| 0 | 1811 | A320 | 1 |
| 1 | 1812 | A320 | 1 |
| 2 | 1813 | B777 | 1 |
| 3 | 1814 | B737 | 1 |
| 4 | 1815 | B737 | 1 |

0s completed at 21:20
Flight.csv
8 / 8
Show all

ProbabilityDisb_1.ipynb - Colab

colab.research.google.com/drive/1JZ-TbmRIY8mmD018C_grjyQHluBGNQxZ#scrollTo=dxAtc_fhXvov

+ Code + Text

RAM Disk

[10]

```
RangeIndex: 1000 entries, 0 to 999
Data columns (total 3 columns):
 #   Column        Non-Null Count  Dtype
---  ------        --------------  -----
 0   Passenger_ID  1000 non-null   int64
 1   Flight_ID     1000 non-null   object
 2   Arrived       1000 non-null   int64
dtypes: int64(2), object(1)
memory usage: 23.6+ KB
```

A _ _ _

```
flights.head()
```

| | Passenger_ID | Flight_ID | Arrived |
|---|---|---|---|
| 0 | 1811 | A320 | 1 |
| 1 | 1812 | A320 | 1 |
| 2 | 1813 | B777 | 1 |
| 3 | 1814 | B737 | 1 |
| 4 | 1815 | B737 | 1 |

✓ 1000 – 9999 ✓

x = 10; programming

stats:

random - variables

arrived = 1 or 0 → binary categorical

✓ passenger-ID = r.v → numeric

1000 – 9999

0s completed at 21:20

Flight.csv

Show all

random-var

↳ categorical / discrete → 1 or 0 (arrived)

↳ p-1D : numeric → a to b

↳ real-valued : e.g : Temp @ a loc on a day

marks

heights / weights

Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:32

INTUITION

{ Probability - distribution

[discrete r.v]

r.v

A arrived $\in \{0, 1\}$

likelihood

$p(A=0) = 0.1$ $\rightarrow 1-0.9$ (let's)

$p(A=1) = 0.9$

11 / 11

Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:34

Q

Fair Dice of 6 sides

r.v → represent outcome of an expt

r.v : $X = \{1, 2, 3, 4, 5, 6\}$

* prob. dist of X

$p(X=1) = \frac{1}{6} = p(X=2) \cdots = p(X=6)$

12 / 12

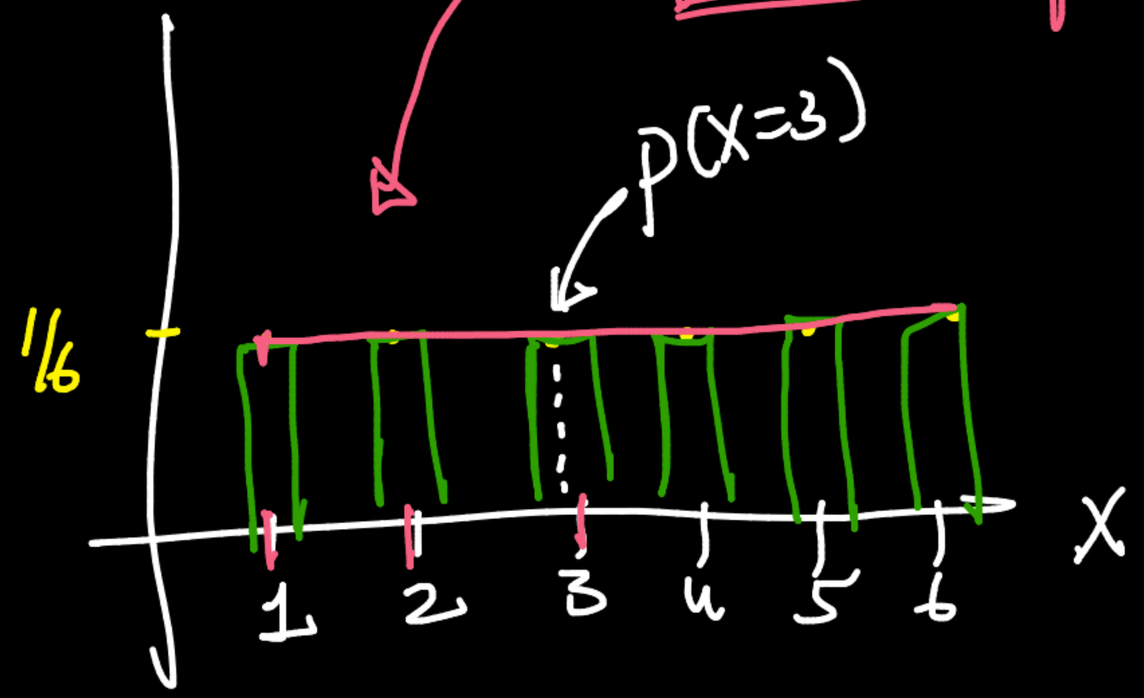
Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:40

dice throwing

Uniform dist of a discrete r.v

PMF

prob. mass fn




$$\left\{ \begin{array}{l} P(X=i) \\ \quad = \text{same} \end{array} \right.$$

i is in $\{1,2,3,4,5,6\}$

13 / 15

Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:38

Q

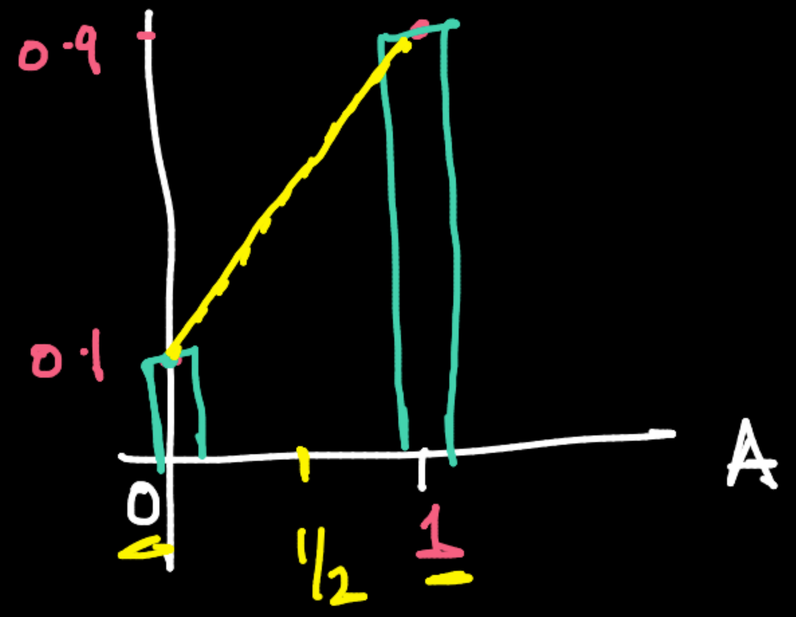
arrived

$A \in \{0, 1\}$

$P(A=0) = 0.1$

$P(A=1) = 0.9$

PMF → discrete r.v




mis understanding

14 / 14

Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:40

Terminology → overwhelming

↳ contextualize

15 / 15

Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:40

16 / 16

Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:41

dice-throwing

$X$ : discrete r.v

↳ distr: Uniform

$X$ is a Uniformly distr discrete r.v

17 / 17

Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:43

Q

$A = \{0, 1\}$ 0.1 0.9

Is this a uniform r.v? → NO ✓

Q

fair coin

$C = \{H, T\}$

↳ Uniform r.v? → Yes ✓

Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:43

PMF: $P(X = i)$

discrete

any of the possible values

19 / 19

Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:45

**CDF:** Cumultative disb fn: (discrete r.v)

Fair dice:

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

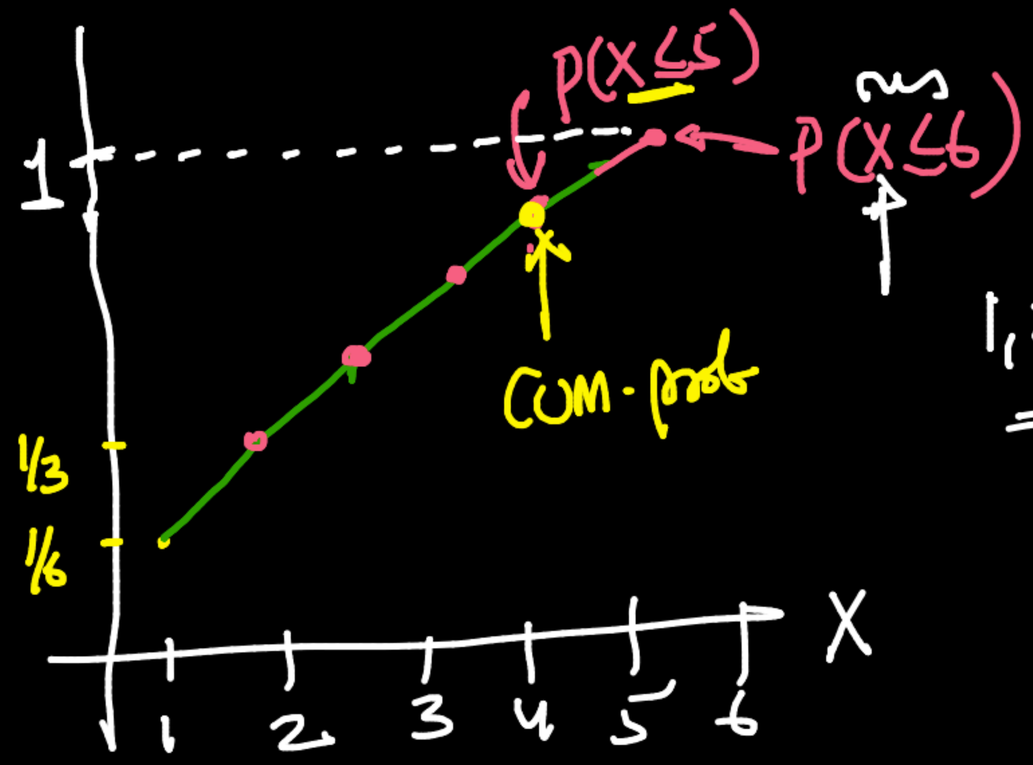
$$P(X \leq 3) = P(X=1) + P(X=2) + P(X=3)$$

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

20 / 20

Cumulative CDF of X:

- non-decr curve
- max-val = 1

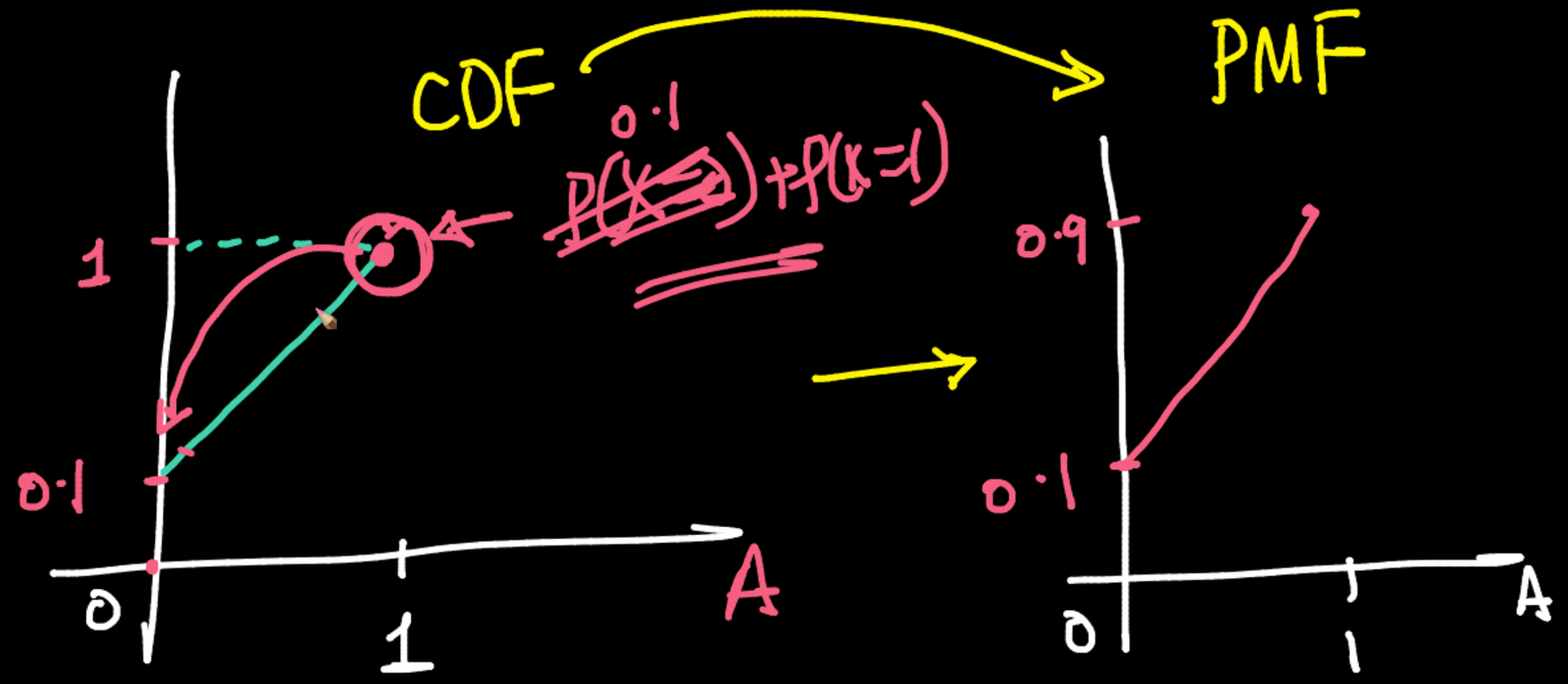



1, 2, 3, 4, 5, 6

$(i, P(X \le i))$

arrived:
A

CDF:




$P(X \le 0) = 0.1$

$P(X=1)$

Q CDF of A

↳ find the PMF of A

Chrome File Edit View History Bookmarks Profiles Tab Window Help 21:53

✓ $\left\{ A: \overbrace{\text{discrete}} \right.$ → numeric

↳ $\leq$

A: Yes, No

$P(A \leq NO) = ?$

COIN: H & T

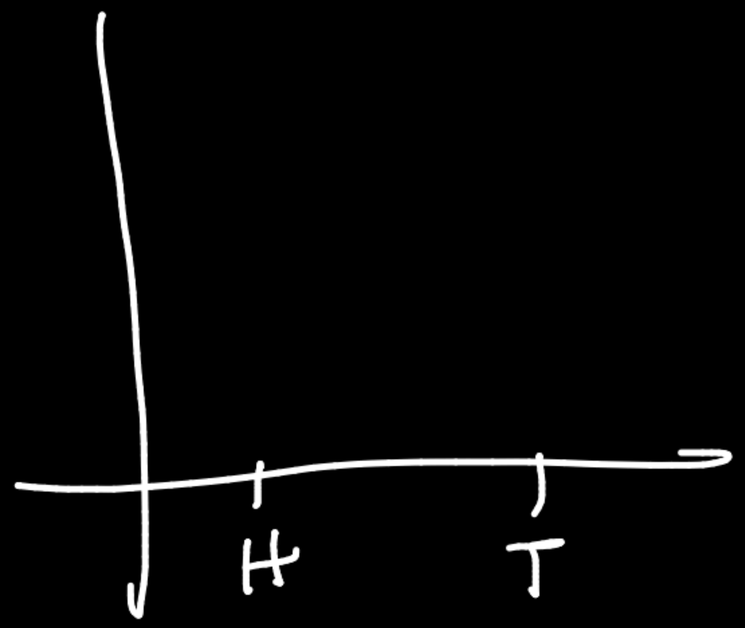
↓ ↓

1 0

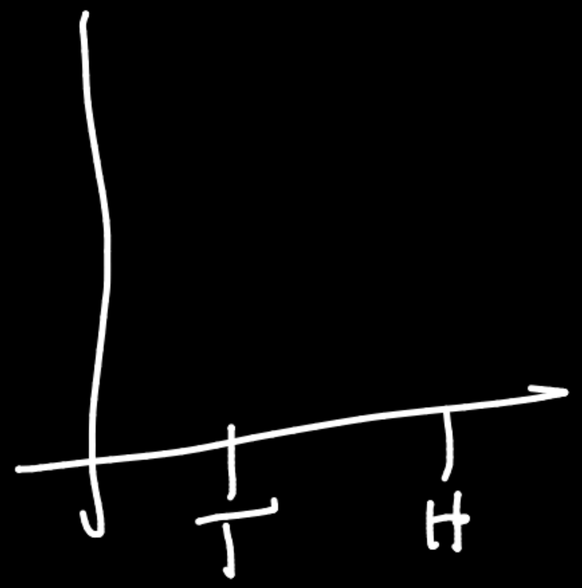
$P(X \leq 1)$

23 / 23

Coin toss:




(or)




$$C: \begin{cases} 0 & \text{if } H \\ 1 & \text{if } T \end{cases}$$

Q

ordering

(ordinal)

$A^+$ : 5%

A : 20%

B : 30%

C : 20%

D : 5%

G : discrete r.v (ordinal)

PMF & CDF

$D < C < B < A < A^+$

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:01

PDF

Continous r.v

real/numeric r.v

X: numeric values

heights

H : heights of learners

Not integer

real-values

150.0 — 190.0 cm

$P(H = 180.0\ cm) =$ close to zero

180.0102345 . . .

INTUITION

26 / 28

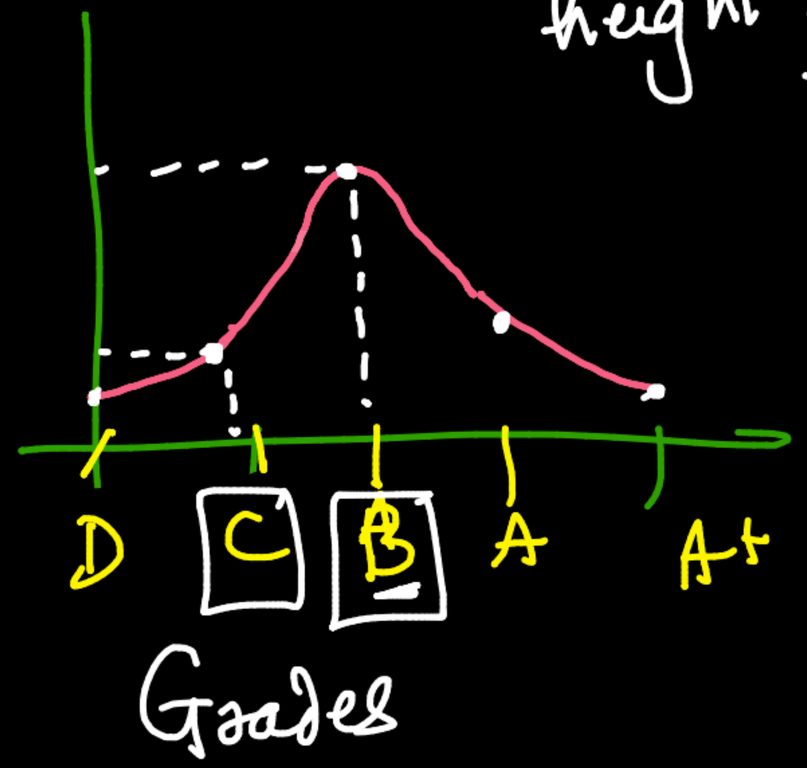
discrete r.V

**PMF:**

$P(X=i)$

not used for continous r.V

(*) height : likelihood



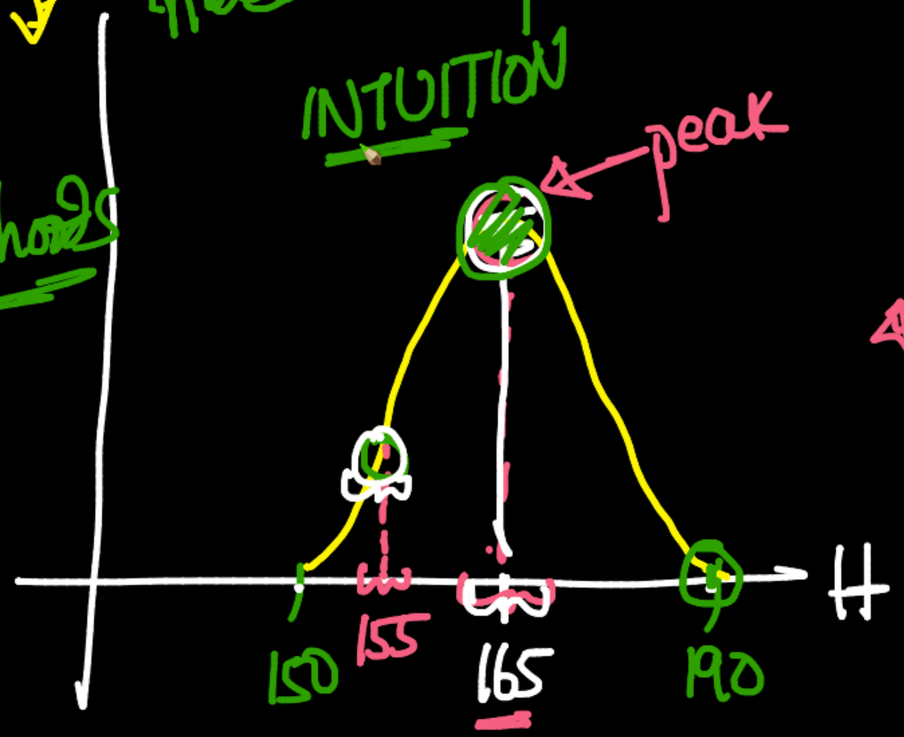
PDF: density

continous r.v

need not represent prob

INTUITION

likelihoods




$f(H) = \cdots$

height ∝ likelihood of the curve

discrete r.v

↳ PMF

Data

Chrome File Edit View History Bookmarks Profiles Tab Window Help
28 / 29

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:12

# Task!

Data:

| | |
|---|---|
| 1 → | 155.2 |
| 2 → | 160.12 |
| 3 → | ⋮ |
| ⋮ | ⋮ |
| 100 → | 181.2 |

→ PDF

165

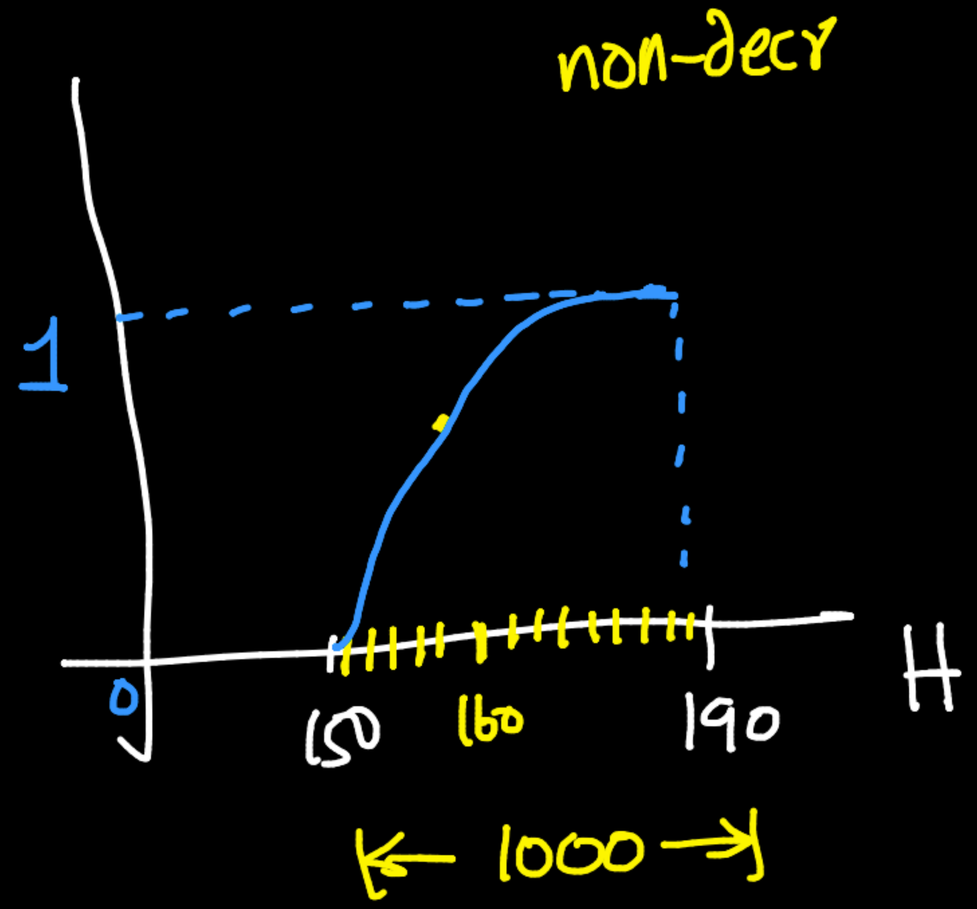
$\frac{33}{100}$

Problem
↓
concepts

29 / 30

discrete r.v → PMF → CDF

continous r.v → $\boxed{\text{PDF}}$ → CDF

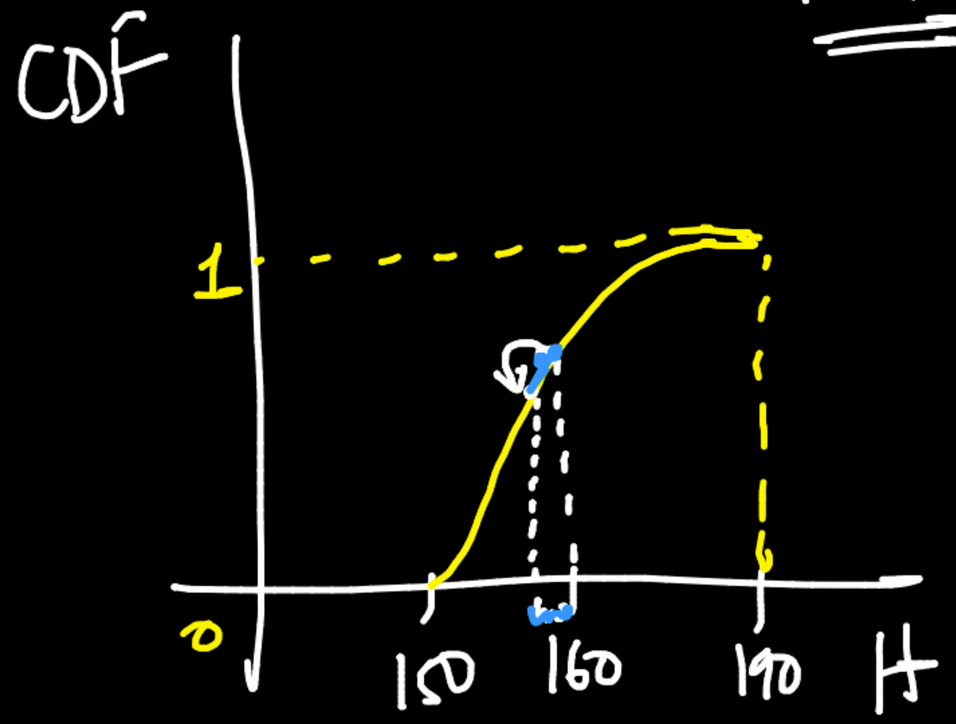


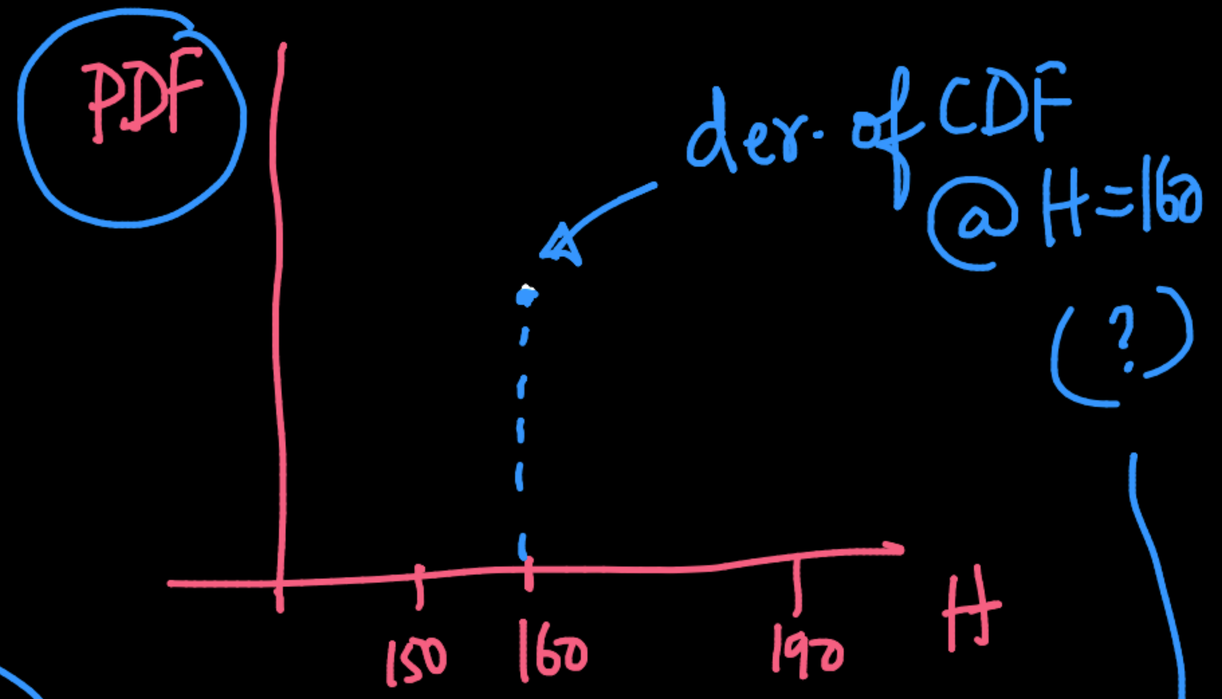

$P(H \leq 160)$

$P(H \leq 190)$

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:19

# Math




159.9999 ← 160

Calculus

31 / 31

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:19

$$\checkmark \begin{cases} \text{PMF}, \ \text{CDF} \longleftarrow \text{disc. r.v} \\ \text{PDF}, \ \text{CDF} \longleftarrow \text{cont. r.v} \end{cases}$$

$$\downarrow$$

32 / 32

```
 1   Flight_ID      1000 non-null   object
 2   Arrived        1000 non-null   int64
dtypes: int64(2), object(1)
memory usage: 23.6+ KB
```

Dice: 1, 2, 3, 4, 5, 6

```
[11] flights.head()
```

| | Passenger_ID | Flight_ID | Arrived |
|---|---|---|---|
| 0 | 1811 | A320 | 1 |
| 1 | 1812 | A320 | 1 |
| 2 | 1813 | B777 | 1 |
| 3 | 1814 | B737 | 1 |
| 4 | 1815 | B737 | 1 |

finite # values

P-ID → discrete or continous

Math

Prac

1000 - 9999

large range of numbers

0s completed at 21:20

Flight.csv

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:23
ProbabilityDisb_1.ipynb - Colab
colab.research.google.com/drive/1JZ-TbmRIY8mmD018C_grjyQHluBGNQxZ#scrollTo=bqaqOwLjl6lc
Update
+ Code + Text
RAM
Disk

[11]

| | | | |
|---|---|---|---|
| 1 | 1812 | A320 | 1 |
| 2 | 1813 | B777 | 1 |
| 3 | 1814 | B737 | 1 |
| 4 | 1815 | B737 | 1 |

```
sns.histplot(flights["Passenger_ID"], kde=True)
plt.show()
```

Data → CDF
CDF —der→ PDF
PDF ← smoothed histogram

Cont-r.v

0s completed at 21:20
Flight.csv
34 / 34
Show all

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:23
ProbabilityDisb_1.ipynb - Colab
colab.research.google.com/drive/1JZ-TbmRIY8mmD018C_grjyQHluBGNQxZ#scrollTo=bqaqOwLjl6lc
Update
+ Code + Text
RAM
Disk

[11]

| | | | |
|---|---|---|---|
| 1 | 1812 | A320 | 1 |
| 2 | 1813 | B777 | 1 |
| 3 | 1814 | B737 | 1 |
| 4 | 1815 | B737 | 1 |

```
sns.histplot(flights["Passenger_ID"], kde=True)
plt.show()
```

histogram, KDE

0s completed at 21:20
Flight.csv
35 / 35
Show all

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:26

ProbabilityDisb_1.ipynb - Colab

colab.research.google.com/drive/1JZ-TbmRIY8mmD018C_grjyQHluBGNQxZ#scrollTo=bqaqOwLjl6lc

+ Code + Text

RAM Disk

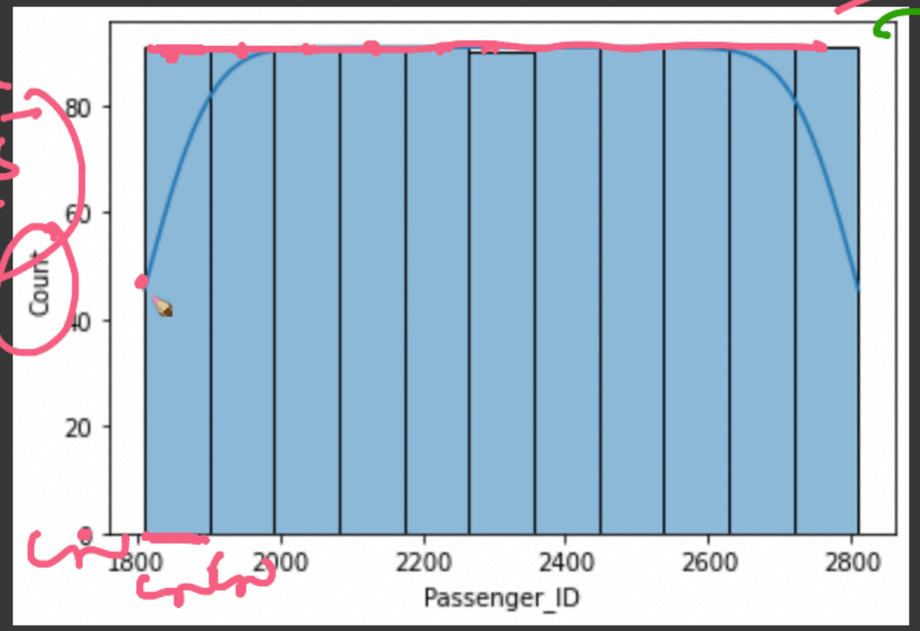
| | | | |
|---|---|---|---|
| 4 | 1815 | B737 | 1 |

→ Smoothing

```
sns.histplot(flights["Passenger_ID"], kde=True)
plt.show()
```

→ PMF

PMF: Prob




min : 1810
max : 2810 } = 1000 obs

Uniform disb

P-ID: discrete r.v

0s completed at 22:23

Flight.csv

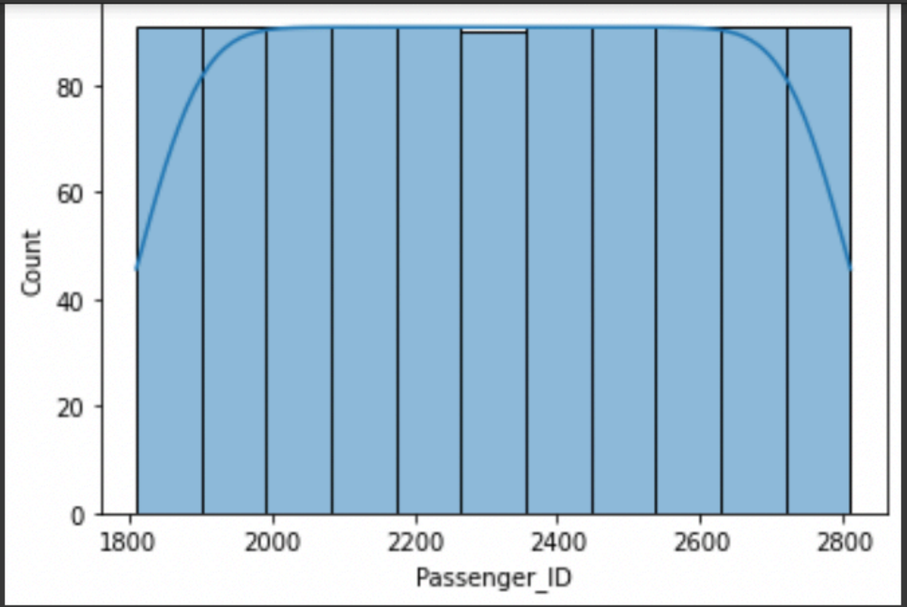
36 / 36

Show all

Task: how much overbooking

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:29
ProbabilityDisb_1.ipynb - Colab
colab.research.google.com/drive/1JZ-TbmRIY8mmD018C_grjyQHluBGNQxZ#scrollTo=8UbrXFk7nQVo
Update
+ Code + Text
RAM
Disk

[12]




$A \to 0 \to \boxed{n_1}$

$A \to 1 \to \boxed{n_2}$

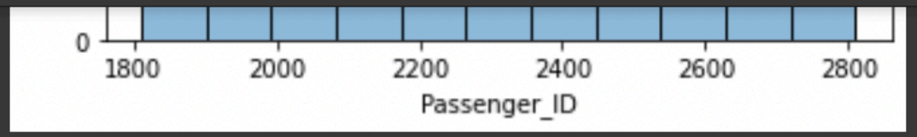
```
flights["Arrived"].value_counts().plot.bar()
plt.show()
```

[ ]

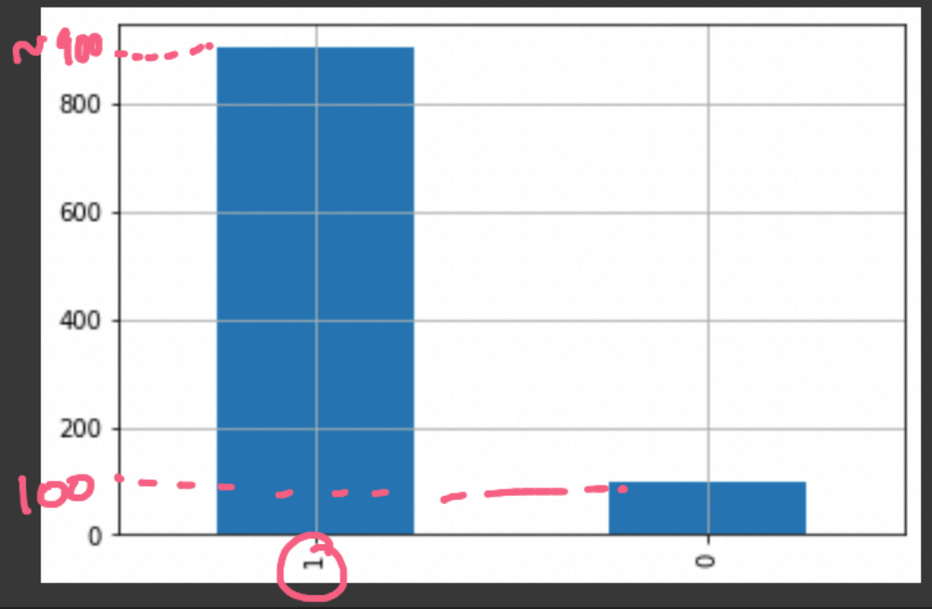
0s completed at 22:23
Flight.csv
38 / 38
Show all

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:30
ProbabilityDisb_1.ipynb - Colab
colab.research.google.com/drive/1JZ-TbmRIY8mmD018C_grjyQHluBGNQxZ#scrollTo=8UbrXFk7nQVo
Update
+ Code + Text
RAM
Disk

[12]




```
flights["Arrived"].value_counts().plot.bar()
plt.grid()
plt.show()
```




$p(A=1) >> p(A=0)$

[ ]

0s completed at 22:29
Flight.csv
39 / 39
Show all

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:30

ProbabilityDisb_1.ipynb - Colab

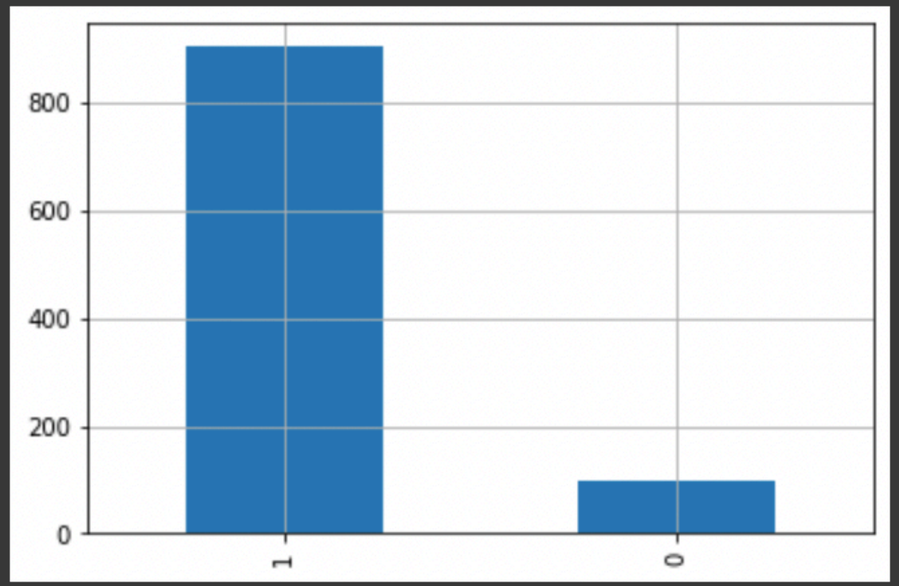
colab.research.google.com/drive/1JZ-TbmRIY8mmD018C_grjyQHluBGNQxZ#scrollTo=JOLsy07TnQNS

+ Code + Text

RAM
Disk

Passenger_ID

[14]
```
flights["Arrived"].value_counts().plot.bar()
plt.grid()
plt.show()
```




0
1

```
showsup_probability = flights['Arrived'].value_counts(normalize=True)[1]
```

0s completed at 22:29

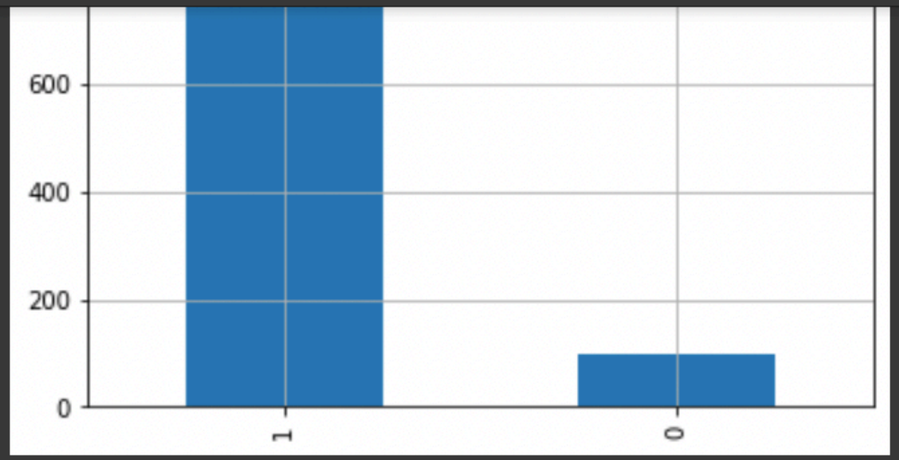
Flight.csv

Show all

40 / 40

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:32
ProbabilityDisb_1.ipynb - Colab
colab.research.google.com/drive/1JZ-TbmRIY8mmD018C_grjyQHluBGNQxZ#scrollTo=-VNFXBP3n-ot
Update
+ Code + Text
RAM
Disk

[14]




[15]
```
showsup_probability = flights['Arrived'].value_counts(normalize=True)[1]
print(showsup_probability)
```

```
0.902
```

```
flights['Arrived'].value_counts()
```

```
1    902
0     98
Name: Arrived, dtype: int64
```

$\frac{902}{1000} = P(A=1)$

0s completed at 22:31
Flight.csv
41 / 41
Show all

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:33
ProbabilityDisb_1.ipynb - Colab
colab.research.google.com/drive/1JZ-TbmRIY8mmD018C_grjyQHluBGNQxZ#scrollTo=-VNFXBP3n-ot
Update
+ Code + Text
RAM
Disk

[15]
```
showsup_probability = flights['Arrived'].value_counts(normalize=True)[1]
print(showsup_probability)
```

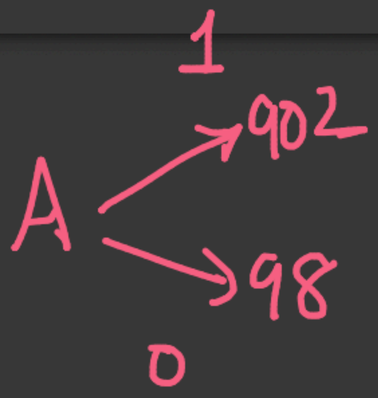
```
0.902
```

$P(A=1) = 0.902$

$P(A=0) = 0.098$

```
flights['Arrived'].value_counts()
```

```
1    902
0     98
Name: Arrived, dtype: int64
```




0s completed at 22:31
Flight.csv
42 / 42
Show all

$A=1$

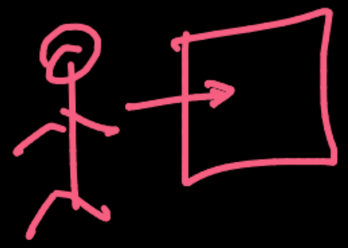
$P(\text{shows-up}) = 0.9$ ← on an average across all planes

$P(\text{not shows up}) = 0.1$

$A=0$

100 ← #tickets

Cap: $\boxed{100}$



overbook: $x$ tickets

90% of $(100+x)$ $\leq 100$ ← Cap

avg-book

overbooking

90% of $(100 + \boxed{x}) \leq 100$ ← cap

$0.9(100+x) \leq 100$

$\boxed{x : 111}$

5000 Rs

if 90% of cust arrive

$x = 10$

(e.g)

total # tickets = 110

90% of 110 = 99 $\leq 100$

(on an avg)

100 → 5L

110 → 5.5L

Penally: → 10,000 Rs for 1 passenger

100 → 500,000

bad-case: 100 + 10 ← overbooked

{ flight-day: 101 people

1 extra

Math:

550,000
−10,000

540,000

Chrome File Edit View History Bookmarks Profiles Tab Window Help 22:46

random

passenger arrives @ airport

A

0 (0.1) ← $1-p=q$

1 (0.9) p

not Uniform r.V

param of the dist

$p=0.9$

PMF; CDF ....

→ takes 2 values

— mutually excl & exh

→ Bernoulli r.V

46 / 46

$\cancel{n!}$

$^nC_1$

$P(\text{Customer arrives}) = 0.9$

$p \rightarrow 0.9 \quad 0.9 \quad 0.9 \quad \cdots \quad 0.9$

$\boxed{1} \quad \boxed{2} \quad 3 \quad 4 \quad \cdots \quad 110$

$P(101 \text{ customers come to the airport out of } 110)$

$\downarrow$

1 penalty

$1 \quad 110 \xrightarrow{C} 101$

$= {}^{110}C_{101} (0.9)^{101} (0.1)^{9}$

$$P(X \text{ and } Y) = P(X) \cdot P(Y)$$

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:00

{ Arrivals → 0, 1 $p = 0.9$

#cust @ the airport

(let)

$X: \; n = 110; \; p = 0.9$

$P(X = 101 \mid X: \text{Bin r-v } p = 0.9, n = 110) = {}^{110}C_{101} \, 0.9^{101} \, 0.1^{9}$

$\sim 11\%$

49 / 51

arrival

A → 0 (0.1)

A → 1 $p = 0.9$

Success

→ Bernoulli r.v

Coin toss → H ($p = 0.5$) ← Success

→ T

Binomial r.v →

$n$ (100) coin toss/trails

$s$ (10) successes

$p$ : prob. of success

X (counts)

$$P(X = s \mid X \text{ is Bin r.v } n = 100, s = 10, p = 0.5) = {}^{n}C_{s}\, p^{s}(1-p)^{n-s}$$

50 / 51

$$P\left(X = 101 \,\middle|\, \begin{array}{c} n = 110;\ p = 0.9 \\ X \sim Bin(\ ) \end{array}\right) = 11\%$$

110 tickets $\Rightarrow$ 10 overbookings

✓ $\left\{ 11\% \text{ chance} \right.$ & penalty of 10,000 ₹

$100 + \boxed{10}$

$110 \longrightarrow$

$X \sim Bin(n=110;\ p=0.9)$

$P(X=101) \longrightarrow 11\% \longrightarrow$

Penalty

$10{,}000 \times \overbrace{0.11}^{} \times 1$

$= 1100$

Penalty: 10,000 dollars

$\sim$ 100 flights $\longrightarrow \sim 11$ flights

1 extra passenger

over

$100 + \boxed{10}$

$110$

$X \sim Bin(n=110;\ p=0.9)$

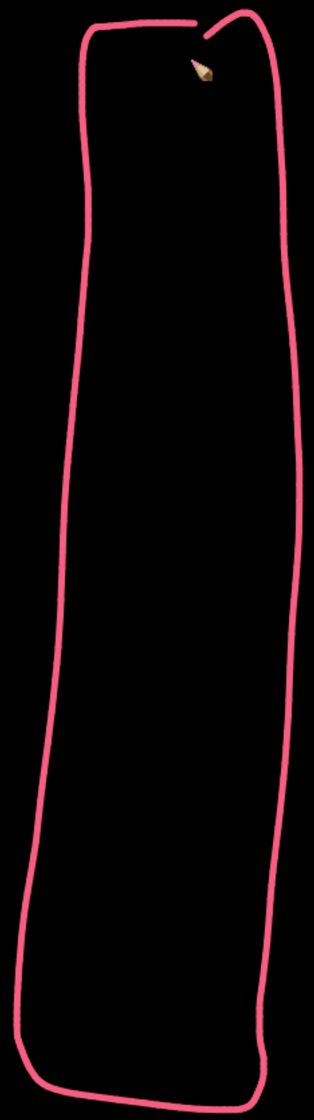
$P(X=102) = {}^{110}C_{102}\ 0.9^{102}\ 0.1^{8} = 8.8\%$

2 penalty

Penalty: $2 \times 10{,}000 \times 0.088 = \boxed{1760}$ Rs

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:08

110 → 101 → 102 → 103 → 104 → ⋮ → 110



54 / 54

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:08

Code:

55 / 55

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:09
ProbabilityDisb_1.ipynb - Colab
pandas.Series.value_counts
colab.research.google.com/drive/1JZ-TbmRIY8mmD018C_grjyQHluBGNQxZ#scrollTo=ih4xtB0CwrJX
Update
+ Code + Text
RAM
Disk

```
0      98
Name: Arrived, dtype: int64
```

```python
import math

PENALTY = 10000
def comb(n, r):

  num1 = math.factorial(n)
  num2 = math.factorial(r)
  num3 = math.factorial(n-r)

  return num1/(num2*num3)
```

$$^{n}C_{r} = \frac{n!}{r!\,(n-r)!}$$

0s completed at 22:31
Flight.csv
56 / 56
Show all

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:11
ProbabilityDisb_1.ipynb - Colab
pandas.Series.value_counts
colab.research.google.com/drive/1JZ-TbmRIY8mmD018C_grjyQHluBGNQxZ#scrollTo=3hVFUORFw5CF
Update
+ Code + Text
RAM
Disk

Simulation

100, 101, 102, - - 150

500,000 5...

[18]
```python
    num1 = math.factorial(n)
    num2 = math.factorial(r)
    num3 = math.factorial(n-r)

    return num1/(num2*num3)
```

110

```python
def calculate_expected_penalty(ticket_sold):

    total_penalty = 0.0

    for i in range(1, ticket_sold - 100+1):
        ##pmf for k successes, n trials, p=success probab
        prob = stats.binom.pmf(k=100+i,n=ticket_sold,p=showsup_probability)
        penalty = prob*PENALTY*i
        total_penalty += penalty

    return total_penalty
```

0s completed at 23:10
Flight.csv
57 / 57
Show all

Chrome File Edit View History Bookmarks Profiles Tab Window Help
ProbabilityDisb_1.ipynb - Colab
pandas.Series.value_counts
colab.research.google.com/drive/1JZ-TbmRIY8mmD018C_grjyQHluBGNQxZ#scrollTo=3hVFUORFw5CF
Update
+ Code + Text
RAM
Disk

[18]

```
num1 = math.factorial(n)
num2 = math.factorial(r)
num3 = math.factorial(n-r)

return num1/(num2*num3)
```

$\boxed{100}$ → 101, 102, 103, 104, 105

1pen 2pen !!

105

```
def calculate_expected_penalty(ticket_sold):

    total_penalty = 0.0

    for i in range(1, ticket_sold - 100+1):
        ##pmf for k successes, n trials, p=success probab
        prob = stats.binom.pmf(k=100+i,n=ticket_sold,p=showsup_probability)
        penalty = prob*PENALTY*i
        total_penalty += penalty

    return total_penalty
```

i : 1,2,3,4,5

0s completed at 23:10
Flight.csv
58 / 59
Show all

ProbabilityDisb_1.ipynb - Colab | pandas.Series.value_counts

colab.research.google.com/drive/1JZ-TbmRIY8mmD018C_grjyQHluBGNQxZ#scrollTo=3hVFUORFw5CF

+ Code + Text

RAM Disk

[18]

```
num1 = math.factorial(n)
num2 = math.factorial(r)
num3 = math.factorial(n-r)

return num1/(num2*num3)
```

$n = 105$
$p = 0.9$

101 102 ... 105

105

```
def calculate_expected_penalty(ticket_sold):

    total_penalty = 0.0

    for i in range(1, ticket_sold - 100+1):
        ##pmf for k successes, n trials, p=success probab
        prob = stats.binom.pmf(k=100+i,n=ticket_sold,p=showsup_probability)
        penalty = prob*PENALTY*i
        total_penalty += penalty

    return total_penalty
```

1,2,3,4,5

$P(X=k \mid X \sim Bin(n=105; p=0.9))$

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

0s completed at 23:10

Flight.csv

59 / 60

Show all

```
[18]  num1 = math.factorial(n)
      num2 = math.factorial(r)
      num3 = math.factorial(n-r)

      return num1/(num2*num3)
```

```
def calculate_expected_penalty(ticket_sold):

    total_penalty = 0.0

    for i in range(1, ticket_sold - 100+1):
        ##pmf for k successes, n trials, p=success probab
        prob = stats.binom.pmf(k=100+i,n=ticket_sold,p=showsup_probability)
        penalty = prob*PENALTY*i
        total_penalty += penalty

    return total_penalty
```

101 102 103 104 105

11% 8.8%

110

1 2 3 4 5

10,000 X 0.088 X 2

Chrome File Edit View History Bookmarks Profiles Tab Window Help
ProbabilityDisb_1.ipynb - Colab
pandas.Series.value_counts
colab.research.google.com/drive/1JZ-TbmRIY8mmD018C_grjyQHluBGNQxZ#scrollTo=3hVFUORFw5CF
Update
+ Code + Text
RAM
Disk

[18]

```
num1 = math.factorial(n)
num2 = math.factorial(r)
num3 = math.factorial(n-r)

return num1/(num2*num3)
```

✓ {SIMULATIONS}

```
def calculate_expected_penalty(ticket_sold):

    total_penalty = 0.0

    for i in range(1, ticket_sold - 100+1):
        ##pmf for k successes, n trials, p=success probab
        prob = stats.binom.pmf(k=100+i,n=ticket_sold,p=showsup_probability)
        penalty = prob*PENALTY*i
        total_penalty += penalty

    return total_penalty
```

0s completed at 23:10
Flight.csv
61 / 62
Show all

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:27

$\rightarrow (1, 2, 3, \ldots 10)$

$\overline{\phantom{xx}}$ 110

$1 \rightarrow 11\% \times 10{,}000 \times 1$

$2 \rightarrow 8.8\% \times 10000 \times 2$

3

4

10

62 / 66

Chrome File Edit View History Bookmarks Profiles Tab Window Help
ProbabilityDisb_1.ipynb - Colab
pandas.Series.value_counts
colab.research.google.com/drive/1JZ-TbmRIY8mmD018C_grjyQHluBGNQxZ#scrollTo=3hVFUORFw5CF
Update
+ Code + Text
RAM
Disk

[18]

```python
    num1 = math.factorial(n)
    num2 = math.factorial(r)
    num3 = math.factorial(n-r)

    return num1/(num2*num3)
```

```python
def calculate_expected_penalty(ticket_sold):

    total_penalty = 0.0

    for i in range(1, ticket_sold - 100+1):
        ##pmf for k successes, n trials, p=success probab
        prob = stats.binom.pmf(k=100+i,n=ticket_sold,p=showsup_probability)
        penalty = prob*PENALTY*i
        total_penalty += penalty

    return total_penalty
```

Cap 100

5000x101

Revenue -

101 → exp. penalty -

102 → "

103 → "

let 121

150 → "

0s completed at 23:10
Flight.csv
63 / 64
Show all

Overbook: 10 tickets

110

99

exactly

airport →

101 → 11% → ✓

102 → 8.8% → ✗

⋮

110 → ...

prob

expected pen.

PMF: $P(X=101 \mid X \sim Bin(n=110, p=0.9))$

$X \begin{cases} P(X \geq 101 \mid X \sim Bin(n=110, p=0.9)) \\ P(X \leq 110 \mid X \sim Bin(n=110, p=0.9)) \end{cases}$

Chrome File Edit View History Bookmarks Profiles Tab Window Help

23:26

66 / 66

```
    for i in range(1, ticket_sold - 100+1):
        ##pmf for k successes, n trials, p=success probab
        prob = stats.binom.pmf(k=100+i,n=ticket_sold,p=showsup_probability)
        penalty = prob*PENALTY*i
        total_penalty += penalty

    return total_penalty
```

```
for i in range(100, 150):
    sales = 5000*i
    penalty  = calculate_expected_penalty(i)
    profit = (sales - penalty)
    print("Overbook seats:",i, ", Profit made:",round(profit))
```

rev/sales — exp.penalty

100 →

101 →

⋮

150 →

```python
for i in range(100, 150):
    sales = 5000*i
    penalty  = calculate_expected_penalty(i)
    netsales = (sales - penalty)
    print("Total seats:",i, ", Net Sales :",round(netsales))
```

```
Total seats: 100 , Net Sales : 500000
Total seats: 101 , Net Sales : 505000
Total seats: 102 , Net Sales : 509996
Total seats: 103 , Net Sales : 514979
Total seats: 104 , Net Sales : 519913
Total seats: 105 , Net Sales : 524725
Total seats: 106 , Net Sales : 529288
Total seats: 107 , Net Sales : 533425
Total seats: 108 , Net Sales : 536929
Total seats: 109 , Net Sales : 539603
Total seats: 110 , Net Sales : 541302
Total seats: 111 , Net Sales : 541959
Total seats: 112 , Net Sales : 541595
Total seats: 113 , Net Sales : 540305
Total seats: 114 , Net Sales : 538233
Total seats: 115 , Net Sales : 535544
Total seats: 116 , Net Sales : 532393
Total seats: 117 , Net Sales : 528919
Total seats: 118 , Net Sales : 525227
```

0s completed at 23:30

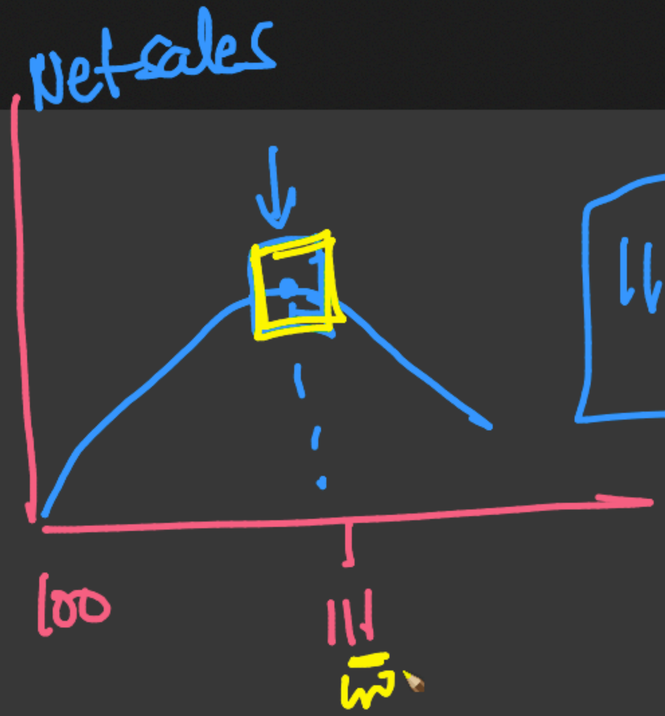
```
for i in range(100, 150):
    sales = 5000*i
    penalty  = calculate_expected_penalty(i)
    netsales = (sales - penalty)
    print("Total seats:",i, ", Net Sales :",round(netsales))
```

```
Total seats: 100 , Net Sales : 500000
Total seats: 101 , Net Sales : 505000
Total seats: 102 , Net Sales : 509996
Total seats: 103 , Net Sales : 514979
Total seats: 104 , Net Sales : 519913
Total seats: 105 , Net Sales : 524725
Total seats: 106 , Net Sales : 529288
Total seats: 107 , Net Sales : 533425
Total seats: 108 , Net Sales : 536929
Total seats: 109 , Net Sales : 539603
Total seats: 110 , Net Sales : 541302
Total seats: 111 , Net Sales : 541959
Total seats: 112 , Net Sales : 541595
Total seats: 113 , Net Sales : 540305
Total seats: 114 , Net Sales : 538233
Total seats: 115 , Net Sales : 535544
Total seats: 116 , Net Sales : 532393
Total seats: 117 , Net Sales : 528919
Total seats: 118 , Net Sales : 525227
```

incr

decr

Simulation




11 ticketsi exlsau

0s completed at 23:30

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:35

# Real-world:

$p = 0.9$ overall

Simulation → day & time

specific models

{ Sat Morning $p = 0.95$

Christmas eve $p = 0.999$

Real-world problem → PMF → Bernoulli, Binomial r.v.

CDF, PDF

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:37

$$A \begin{cases} \to 0 \\ \to 1 \end{cases} \quad p = 0.9$$

$$\text{Bin}(n,p) \sim X : \underline{\underline{PMF}} \longrightarrow \underline{\underline{+\text{code}}} \ \& \ \text{Simulation} \ \checkmark$$

71 / 71

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:42

2-values

A → 0, 1 (success)

Bernoulli r.V (discrete)

X: counting # successes of a Bernoulli trail → Binomial r.V discrete

1, 2, . . .

$X \sim Bin(n=1, p)$

72 / 73

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:41

[Multinomial r.V]

Dice → discrete Uniform disb r.V

NOT bernoulli

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:45

Q

C → Yes = 1, No = 0

$p = 0.15$

Binary

discrete; Bernouli; not uniform

A: Out of $\overbrace{1000}^{n2}$ people; count # people who got covid

$n = 1000$

0, 1, 2, ....

discrete
NOT binary
Binomial

74 / 74

Covid test result of a freq-flyer

Q

A: → Y, N  $p = 0.5$  binary,

Count the number of times a person gets infected with covid  10 times

V → $0.5 \times 0.5 \cdots \times 0.5$

Varma

A: → $p = 0.3$

Binomial r.v

(n times)
100

K times
20

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:52
76 / 76

Chrome File Edit View History Bookmarks Profiles Tab Window Help 23:53

ProbabilityDisb_1.ipynb - Colab | pandas.Series.value_counts

colab.research.google.com/drive/1JZ-TbmRI

+ Code + Text

```
[18]  return num1/(nu
```

```
def calculate_exp

    total_penalty

    for i in rang
        ##pmf for
        prob = st
        penalty =
        total_pen

    return total_
```

```
[21] for i in range(10
        sales = 5000*
        penalty  = ca
        netsales = (s
        print("Total
```

```
Total seats: 100
Total seats: 101
```

Flight.csv

Live | DSML Advanced : Probab

scaler.com/meetings/i/dsml-advanced-probability-distributions-1-2/live

# You have left the meeting

## We get frequent requests for your notes!

Notes written by you helps in understanding the topic better. You can upload the notes in two simple steps mentioned below

1. **Scan the QR code with your iPad**
   Scanner should be present in the top menu on your iPad
2. **Upload Notes on the generated link**
   All notes uploaded will be visible in the saved version of this session

OR

**Drag and drop files** or **click here to upload**

Files Uploaded from your computer appear here

77 / 77

Show all