9th December, 2022

DSML : CC Maths

Probability 8 - Distributions - 3.

Recap:
(a) Probability theory.
(b) Bayes' theorem.
(c) Combinatorics.
(d) Descriptive statistics.
(e) Binomial Distribution.
(f) Gaussian Distribution.

Today:
(a) Poisson Distribution
(b) Geometric Distribution.
(c) Problem solving.

Class starts @ 9:05 p.m.

Poisson : Time or space bound activities.

Football : The average number of goals in a 90 min match is 2.5.

How many goals in 30? $\rightarrow$ 2.5/3.

Rate : $\boxed{2.5 \,/\, 90 \text{ mis}}$

1.25 / 45 mins.

$\simeq$ 0.83 / 30 mins.

Customers entering a store : $\boxed{100 / \text{day}}$.

$\approx$ 4 in 1 hour?

Rate : 100 / day.

Support center phone calls : $\boxed{100 \text{ calls / hour}}$.

$100/60$. in 1 minute

→ Farmer : $\boxed{100 \text{ trees in 1 acre.}}$

can there be more than 60 trees in 1/2 acre?

→ Hospital : $\approx$ $\boxed{5 \text{ patiens / hour}}$, emergency.

→ What is the probability that we will have more than 10 in the next hour?

→ Typos : $\approx$ $\boxed{3 \text{ typos / page}}$.

$P(X = 0)$

Rate : expected number of events per interval.

The interval is typically time; but can be space or the num. pages, no. of goals etc.

λ → Average occurrence

μ.

Poisson Distribution :

1] **Counting # occurrences.**
The experiment should count events in some interval.

2] **Independence :**
The occurrence of one event does not affect the probability that another will occur.

3] **Rate :** → The rate is fixed / constant.
The average rate at which events occur is independent of any occurrences.

4] **No Simultaneous Events :**
No events occur simultaneously.

Poisson

A city sees 3 accidents / day.

P(X = 5) tomorrow?

3 / day.

Let x denote the accidents tomorrow.

x is Poisson distributed with $\lambda / \mu = 3$.

$E[x] = 3. \rightarrow \mu = 3$

from scipy.stats import poisson.

$P(x = 5) = 0.1008.$

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

X : # of typos in a page. The average number of typos / page is 3.

What is the probability of "at most" 1 typo?

$P(X=0) + P(X=1) = 0.199.$

$$\underbrace{P(X=0)}\ \to\ \frac{\lambda^k \cdot e^{-\lambda}}{k!} = \frac{\lambda^0 \cdot e^{-\lambda}}{0!}$$

$$= e^{-3} = \underline{0.049.}$$

$P(X \leq 1) \to$ cdf.

A shop is open for 8 hours. The average number of customers is 74.

Q1] Expected customers in 2 hrs?

18.5 :

Q2] What is the probability that in 2 hours, at most 15 customers?

18.5 / 2 hours.

Quiz time!

Quiz Ended!

$P(X \le 15)$

cdf.

A shop is open for 8 hours. The average number of customers is 74 - assume Poisson distributed. What is the probability that in 2 hours, there will be at most 15 customers?

24 users have participated

A	0.16	0%
B	0.185	4%
C	0.203	13%
D	0.249	75%
E	0.295	8%
F	0.31	0%

$8 - 74$

$2 - ?$

$= \quad x = \frac{74 \times 2}{8} = \cancel{18.5}$

2hours

A shop is open for 8 hours. The average number of customers is 74.

Q] Prob. that in 2 hours, there will be at least 7 customers?

18.5/2 hours.

$$P[X \geq 7] = 1 - P[X < 7] \quad \text{cdf}$$

A shop is open for 8 hours. The average number of customers in 8 hours is 74
What is the probability that in 2 hours, there will be at least 7 customers?

40 users have participated

A	0.87	5%
B	0.91	18%
C	0.95	2%
D ✓	0.99	75%

$$1 - P[X \leq 6].$$

You receive 240 messages / hour on average.

Q] What is the expected no. of messages in 30s?

3600s — 240.

30s — ?

$$x = \frac{\cancel{30}^{\,} \times \cancel{240}^{\,2}}{\cancel{3600}_{\,\cancel{12}}} = 2.$$

Q]. 1 message in 30s.

Q] Prob. that there are no messages in 15 seconds.

Q] Prob. that there are 3 messages in 20s.

$\mu / \lambda = 2$

$P[X = 1]$.

Quiz time!

Quiz Ended!

You receive 240 messages per hour on average - assume Poisson distributed. What is the probability of one message arriving over a 30 second time interval?

38 users have participated

A	0.15	3%
B	0.21	3%
C	0.27	63%
D	0.36	3%
E	0.39	3%
F	0.4	26%

$P[X = k]$

pmf.

$(1-)\ P[X \leq k]$

CDF.

$P(X > k)$

Quiz time!

Quiz Ended!

You receive 240 messages per hour on average - assume Poisson distributed. What is the probability that there are no messages in 15 seconds?

41 users have participated

A	0.251	10%
B ✓	0.367	85%
C	0.412	5%
D	0.428	0%

30 — 2

15 — ?

$\frac{15}{30} \times 2 = 1.$

$P[X = 0]$

Rate = 1.

$P[X = 3]$ $\lambda = ?/20s.$

$3600 - 240$

$20 - ?$

$$x = \frac{20 \times 240}{3600}$$

$= 1.33.$

$P[X = 3]$

$k = 3$, $mu = 1.33.$

Geometric Distribution: Applied when we stop after the 1st success.

eg: Suppose we throw a dice till we get 6.

Q] What is the probability that we have to throw it 4 times?

from scipy.stats import geom.

$P[X = 4]$

$\underbrace{P[6]}_{p} = \frac{1}{6}$ $\qquad$ $P[6'] = \frac{5}{6}.$

$$P[X = k] = \underbrace{(1-p)^{k-1}}\cdot p$$

$p \rightarrow$ prob. success

$1-p \rightarrow$ prob. failure.

eg: suppose we throw a dice till we get 6.

Q1] $P[X = 4]$

Q2] Prob. we have to try at most 3 times.

$P[X \leq 3]$. $1 - P[X \leq 2]$.

Q3] Prob. that we have to try at least 3 times. $P[X > 2]$.

$1 - cdf(k=2, p=1/6)$.

"At least 3"

$$P[X=3] + P[X=4] + P[X=5]$$

$$\dots + \infty.$$

$$\sum_{k=3}^{\infty} P[X=k].$$

Q] I am playing a game where the probability of winning a prize is 0.7.

$p = 0.7.$

$X = 1, 2, 3, \ldots$

Q1] $P[X = 4]$?

Q2] Probability that I don't win in the 1st two attempts?

$P[X > 2].$

$1 - P[X \leq 2].$

cdf.

$P(F) = 0.3.$

$k = 2.$

$$\underline{0.7 \times 0.3 \quad \times}$$

$P[X > 2].$

$1 - P[X \leq 2]$

$$1 - \left[\underbrace{P[X = 1]}_{p} + \underbrace{P[X = 2]}_{(1-p)p.} \right].$$

What is the expected number of attempts to win the prize?

↓

1.42

↓

geom.expect(args=(0.7,)).

$P[X=3] + P[X=4] + P[X=5] + \dots$

$(1-p)^2$ $(1-p)^2 p$ $(1-p)^3 \cdot p$.

Geometric distribution `geom.pmf(k, p)` $P[X=k] = (1-p)^{k-1}p$

$E[X] = \frac{1}{p}$

I am playing a game where the prob of winning a prize is 0.7

What is the probability that I win the prize on the 4th attempt?

$P[X=4] = (0.3)^3(0.7) = 0.0189$

```
from scipy.stats import geom
```

$P[X=4] =$ `geom.pmf(k=4, p=0.7)` $= 0.0189$

What is the probability that I don't win in the first two attempts

$P[X>2] = 1 - P[X \le 2]$

$P[X>2] =$ `1 - geom.cdf(k=2, p=0.7)` $= 0.09$

What is the expected number of attempts to win the prize

$E[X] = \frac{1}{0.7} = 1.42$

$E[X] =$ `geom.expect(args=(0.7,))` $= 1.42$

$p = 0.7$.

Probability that I don't win in first two attempts?

Q] Are the following valid?

1) Rahul : Probability of failing 2 times?

$0.3 \times 0.3 = 0.09.$

Yes.

2) $p = 0.3.$

$k = 2.$

$P[X = 2]$

$\underline{(1-p)} \cdot \underline{p}$

$(1-p) \to 0.7$, $p \to 0.3$

$0.7 \times 0.3 \to 0.21.$

$p + p(1-p)$ $p = 0.3$.

$0.3 + 0.3 \times 0.7$

$= 0.51$.

pmf → probability mass function.

↳ Discrete Random Variables.

pdf → probability distribution function

↳ Continuous random variables. pdf.

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}.$$

$$\underbrace{P[X=3) + P[X=4) + \cdots\cdots}$$

$$P[X \geq 3]. \qquad 1 - P[X \leq 2]$$

"at least"

cdf (k = 2 , p = 0.7).

$$P[X \leq 3] \rightarrow \text{cdf}.$$

[poisson.pmf(x) for x in range(0, ∞)]

for poisson.

$X = 0, 1, 2, 3, \ldots \infty$.

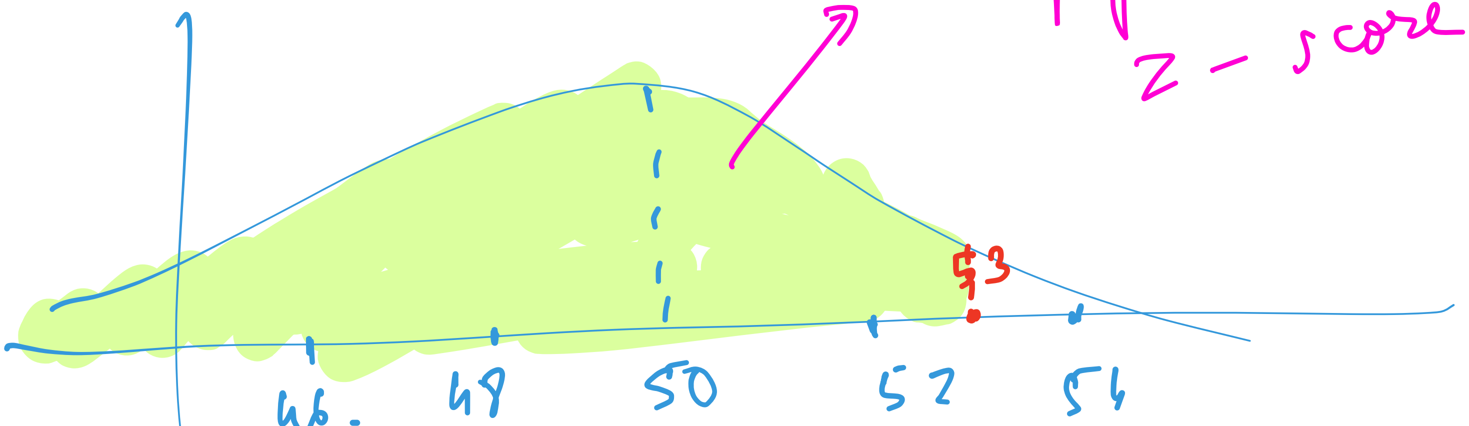
[geom.pmf(x) for x in range(1, ∞)]

pmf.

λ

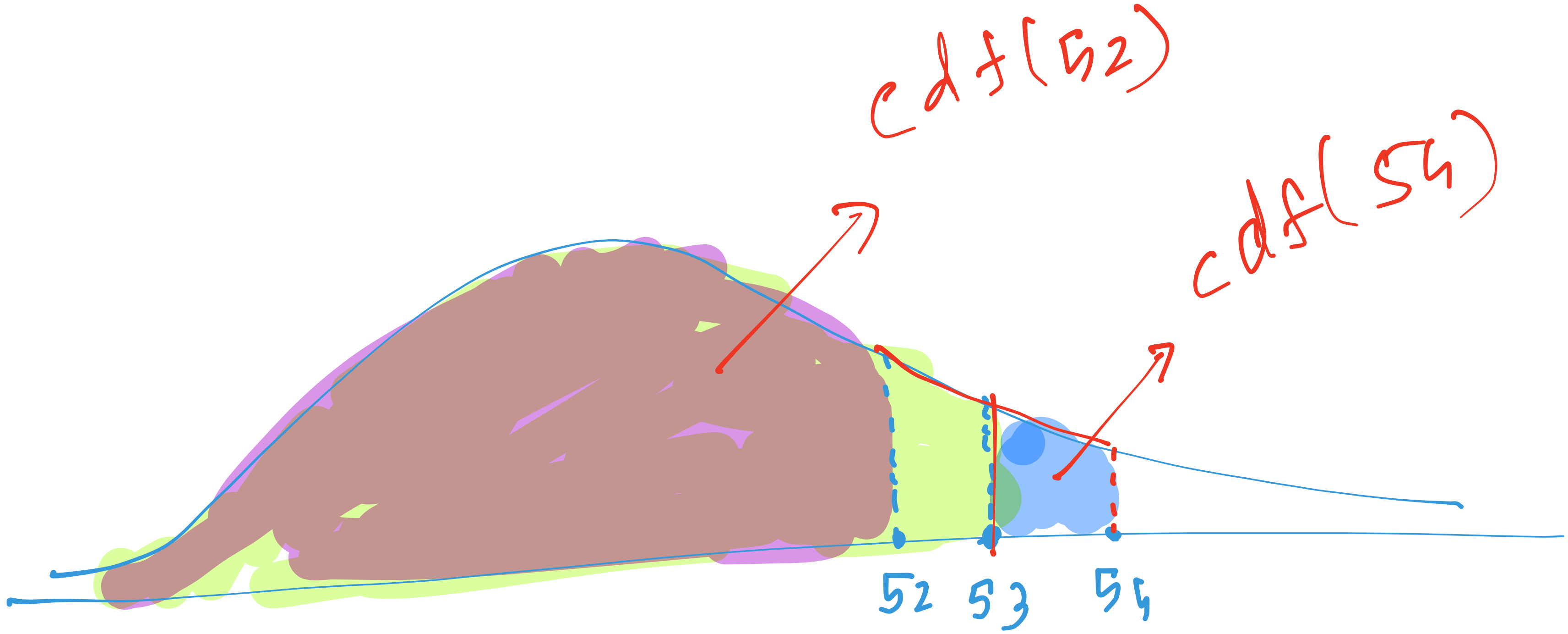
$\mu = 50mm$

$\sigma = 2mm.$

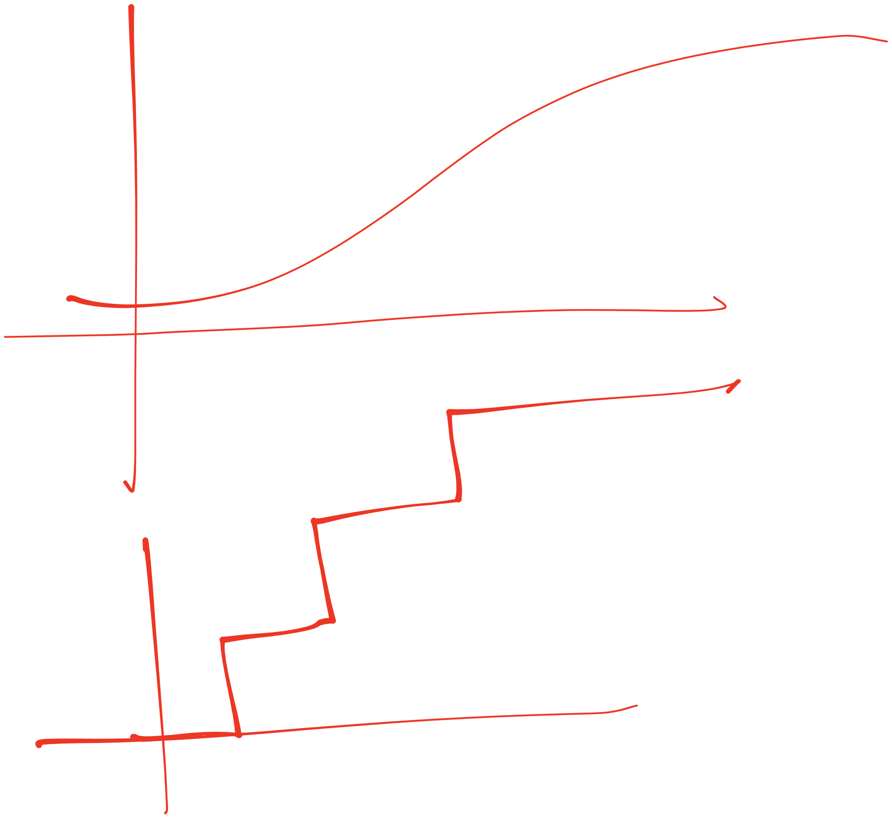


without python or z-score table.

$$\int_{-\infty}^{53} \frac{1}{2\sqrt{2\pi}} e^{-\frac{(x-50)^2}{8}}.$$



$$\frac{cdf(54) + cdf(53)}{2}$$



Cumulative distribution function.