Casino case study

A bag has 3 red and 2 blue balls.

You pick a ball, write its colour, and put it back in the bag. This is done 4 times in total.

If all 4 times, the red ball was drawn, you win Rs 150. In any other case, you lose Rs 10.

Would you play this game?

Let "X" denote the number of red balls when you draw 4 balls with replacement

Here, X is an example of what is called a "Random Variable"

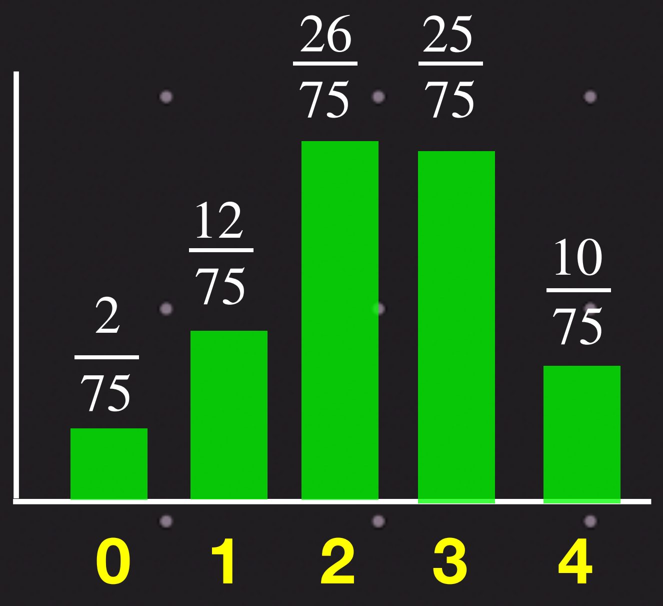
What are all the outcomes?

0 red 1 red 2 red 3 red 4 red

Empirical approach: Estimate probability using data

Data from 75 people

$X = 0$	2 people
$X = 1$	12 people
$X = 2$	26 people
$X = 3$	25 people
$X = 4$	10 people



X	$P[X]$	$E[X]$
0	$\frac{2}{75}$	$(0)\left(\frac{2}{75}\right) +$
1	$\frac{12}{75}$	$(1)\left(\frac{12}{75}\right) +$
2	$\frac{26}{75}$	$(2)\left(\frac{26}{75}\right) +$
3	$\frac{25}{75}$	$(3)\left(\frac{25}{75}\right) +$
4	$\frac{10}{75}$	$(4)\left(\frac{10}{75}\right)$

Expectation of X is the weighted average of the values that X takes, with the weights being the probabilities

$$E[X] = (0)\left(\frac{2}{75}\right) + (1)\left(\frac{12}{75}\right) + (2)\left(\frac{26}{75}\right) + (3)\left(\frac{25}{75}\right) + (4)\left(\frac{10}{75}\right) = 2.38$$

Casino case study

A bag has 3 red and 2 blue balls.

You pick a ball, write its colour, and put it back in the bag. This is done 4 times in total.

If all 4 times, the red ball was drawn, you win Rs 150. In any other case, you lose Rs 10.

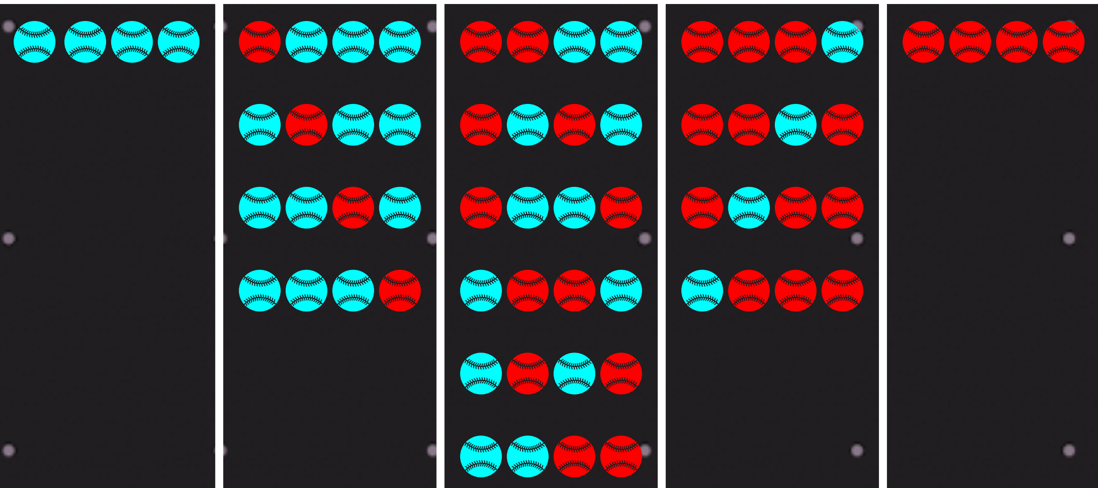
Would you play this game?

Let "X" denote the number of red balls when you draw 4 balls with replacement

Here, X is an example of what is called a "Random Variable"

What are all the outcomes?

0 red	1 red	2 red	3 red	4 red



0 red	1 red	2 red	3 red	4 red
$\frac{2}{5}\frac{2}{5}\frac{2}{5}\frac{2}{5}$	$\frac{2}{5}\frac{2}{5}\frac{2}{5}\frac{3}{5}$	$\frac{2}{5}\frac{2}{5}\frac{3}{5}\frac{3}{5}$	$\frac{2}{5}\frac{3}{5}\frac{3}{5}\frac{3}{5}$	$\frac{3}{5}\frac{3}{5}\frac{3}{5}\frac{3}{5}$
${}^4C_0$	${}^4C_1$	${}^4C_2$	${}^4C_3$	${}^4C_4$

Theoretical approach: Compute probability using rules

X	Number of outcomes	Probability per outcome	$P[X]$	Code
0	${}^4C_0$	$\left(\frac{2}{5}\right)^4$	${}^4C_0\left(\frac{2}{5}\right)^4$	`binom.pmf(k=0, n=4, p=3/5)`
1	${}^4C_1$	$\left(\frac{2}{5}\right)^3\left(\frac{3}{5}\right)^1$	${}^4C_1\left(\frac{2}{5}\right)^3\left(\frac{3}{5}\right)^1$	`binom.pmf(k=1, n=4, p=3/5)`
2	${}^4C_2$	$\left(\frac{2}{5}\right)^2\left(\frac{3}{5}\right)^2$	${}^4C_2\left(\frac{2}{5}\right)^2\left(\frac{3}{5}\right)^2$	`binom.pmf(k=2, n=4, p=3/5)`
3	${}^4C_3$	$\left(\frac{2}{5}\right)^1\left(\frac{3}{5}\right)^3$	${}^4C_3\left(\frac{2}{5}\right)^1\left(\frac{3}{5}\right)^3$	`binom.pmf(k=3, n=4, p=3/5)`
4	${}^4C_4$	$\left(\frac{3}{5}\right)^4$	${}^4C_4\left(\frac{3}{5}\right)^4$	`binom.pmf(k=4, n=4, p=3/5)`

$E[X] = 2.4$

```
binom.expect(args=(4, 3/5))
```

Suppose we toss a coin once every 10 minutes.
The probability of heads is 0.2778

```
binom.pmf(k, n, p)
binom.expect(args=(n, p))
```

What is the probability of getting one heads in 30 minutes?

Number of tosses in 30 minutes: n = 3   p = 0.2778   k = 1

$P[X=1] =$ `binom.pmf(k=1, n=3, p=0.2778)` $= 0.4346$

What is the expected number of heads in 90 minutes?

Number of tosses in 90 minutes: n = 9   p = 0.2778

$$E[X] = n * p = 9 * 0.2778 = 2.5$$

$E[X] =$ `binom.expect(args=(9, 0.2778))` $= 2.5$

Suppose we toss a coin once every minute.
The probability of heads is 0.02778

```
binom.pmf(k, n, p)
binom.expect(args=(n, p))
```

What is the probability of getting one heads in 30 minutes?

Number of tosses in 30 minutes: n = 30   p = 0.02778   k = 1

$P[X=1] =$ `binom.pmf(k=1, n=30, p=0.02778)` $= 0.3681$

What is the expected number of heads in 90 minutes?

Number of tosses in 90 minutes: n = 90   p = 0.02778

$$E[X] = n * p = 90 * 0.02778 = 2.5$$

$E[X] =$ `binom.expect(args=(90, 0.02778))` $= 2.5$

Suppose we toss a coin 10 times every minute.
The probability of heads is 0.002778

```
binom.pmf(k, n, p)
binom.expect(args=(n, p))
```

What is the probability of getting one heads in 30 minutes?

Number of tosses in 30 minutes: n = 300   p = 0.002778   k = 1

$P[X = 1] =$ `binom.pmf(k=1, n=300, p=0.002778)` $= 0.3627$

What is the expected number of heads in 90 minutes?

Number of tosses in 90 minutes: n = 900   p = 0.002778

$$E[X] = n * p = 900 * 0.002778 = 2.5$$

$E[X] =$ `binom.expect(args=(900, 0.002778))` $= 2.5$

Poisson distribution

Typically a time or space bound activity

Football game

Average goals every 90 mins is 2.5

Rate = 2.5 goals / 90 mins

We may be interested in probability of 1 goal in the first half of the match

Rate = 1.25 goals / 45 mins

Customers entering a store

Some 100 customers arrive every day

Rate = 100/day

We may be interested in probability of 10 customers in the next hour

Support centre phone calls

Some 100 calls every hour.

Rate = 100/hour

We may be interested in knowing optimal number of staff

Rate = 1.66/minute

Poisson distribution

Typically a time or space bound activity

Farmers delight

Suppose there are 100 trees every acre of land

Can there be more than 60 trees in half an acre?

Rate = 100 trees / acre

Hospital emergency

Suppose, on average, 5 patients come every hour

What is the probability of more than 10 people next hour?

Rate = 5 patients / hour

Typos

A book might have an average of 3 typos per page

What is the probability of a page having no typos?

Rate = 3/page

Poisson distribution   Typically a time or space bound activity

Rate is the average or expected number of events per interval

This interval is typically time, but can be space, or even “number of pages” etc.

We typically denote Rate = λ

Because it is also the average or expected number, some literature may also use Rate = μ

Poisson distribution   Rules deciding Poisson

Counting

The experiment counts the number of occurrences of an event over an interval

Independence

The occurrence of one event does not affect the probability that a second event will occur

Rate

The average rate at which events occur is independent of any occurrences

No Simultaneous events

No two event occur simultaneously

Poisson distribution

A city sees 3 accidents per day on average.

Find the probability that there will be 5 accidents tomorrow

Rate $\lambda = 3$ per day

Let "X" denote the number of accidents tomorrow

We say "X" is Poisson distributed with rate = 3

E[X] = 3 $\mu = 3$ same as λ

$P[X = 5] =$ `poisson.pmf(k=5, mu=3)`

```
from scipy.stats import poisson
```

$= 0.1008$

$$P[X = 5] = \frac{3^5 e^{-3}}{5!}$$

$$P[X = k] = \frac{\lambda^k e^{-\lambda}}{k!}$$

Poisson distribution

`poisson.pmf(k, mu)` $P[X=k] = \frac{\lambda^k e^{-\lambda}}{k!}$

A city sees 3 accidents per day on average.

Find the probability that there will be at most one accident tomorrow?

$\lambda = 3$

$P[X=0] =$ `poisson.pmf(k=0, mu=3)` $= 0.049$

$$P[X=0] = \frac{3^0 e^{-3}}{0!} = e^{-3} = 0.049$$

$P[X=1] =$ `poisson.pmf(k=1, mu=3)` $= 0.149$

$$P[X=1] = \frac{3^1 e^{-3}}{1!} = 3e^{-3} = 0.149$$

$$P[X \leq 1] = P[X=0] + P[X=1] = 4e^{-3} = 0.199$$

Poisson distribution

`poisson.pmf(k, mu)` $P[X=k] = \frac{\lambda^k e^{-\lambda}}{k!}$

You receive 240 messages per hour on average - assume Poisson distributed.

Q1) What is the average or expected number of messages in 30 seconds?

1 hour (3600 seconds) — 240 messages

30 seconds — ?

$$\frac{30 * 240}{3600} = 2$$

Q2) What is the probability of one message arriving over a 30 second time interval?

If we consider 30 seconds as one unit interval, then λ = 2

$P[X=1] =$ `poisson.pmf(k=1, mu=2)` = 0.27

$$P[X=1] = \frac{(2)^1 e^{(-2)}}{1!} = 0.27$$

Poisson distribution

`poisson.pmf(k, mu)` $P[X = k] = \dfrac{\lambda^k e^{-\lambda}}{k!}$

A shop is open for 8 hours. The average number of customers is 74 - assume Poisson distributed.

Q1) What is the average or expected number of customers in 2 hours?

8 hours 74 customers

2 hours ?

$$\frac{2 * 74}{8} = 18.5$$

Q2) What is the probability that in 2 hours, there will be at most 15 customers?

$P[X \leq 15] =$ `poisson.cdf(k=15, mu=18.5)` $= 0.249$

Q3) What is the probability that in 2 hours, there will be at least 7 customers?

$P[X \geq 7] = 1 - P[X \leq 6] =$ `1 - poisson.cdf(k=6, mu=18.5)` $= 0.99$

Poisson distribution

`poisson.pmf(k, mu)` $P[X = k] = \frac{\lambda^k e^{-\lambda}}{k!}$

Suppose we receive 3 support tickets every 20 days.

Q1) What is the average or expected number of tickets in a day?

20 days		3 tickets
1 day		?

$\frac{3}{20} = 0.15$

Q2) What is the probability that there will not be more than 1 ticket in a day?

$P[X \leq 1] =$ `poisson.cdf(k=1, mu=0.15)` $= 0.989$

Poisson distribution

`poisson.pmf(k, mu)` $P[X = k] = \frac{\lambda^k e^{-\lambda}}{k!}$

Football games have an average of 2.5 goals every 90 mins

Q1) What is the average or expected number of goals in 30 mins?

90 mins	2.5 goals
30 mins	?

$$\frac{30 * 2.5}{90} = 0.8333$$

Q2) What is the probability that there will be exactly 1 goal in 30 mins?

$P[X = 1] =$ `poisson.pmf(k=1, mu=0.8333)` $= 0.362$

Poisson distribution

`poisson.pmf(k, mu)`

$$P[X = k] = \frac{\lambda^k e^{-\lambda}}{k!}$$

There are 80 students in a kinder garden class.

Each one of them has 0.015% chance of forgetting their lunch on any given day.

Q1) What is the average or expected number of students who forgot lunch in the class?

1 student — 0.015

80 students — ?

$$\frac{80 * 0.015}{1} = 1.2$$

Q2) What is the probability that exactly 3 of them will forget their lunch today?

$P[X = 3] =$ `poisson.pmf(k=3, mu=1.2)` $= 0.086$

$$P[X = 3] = \frac{(1.2)^3 e^{-1.2}}{3!} = 0.086$$

Binomial distribution

Here, n = 80, p = 0.015, and k = 3

$$\lambda = np$$

$P[X = 3] =$ `binom.pmf(k=3, n=80, p=0.015)`

Binomial trials "n" is at least 30
Probability of success "p" is at most 0.05

Binomial $E[X] = np$

$$\lambda = np$$