ECS7024 Statistics for Artificial Intelligence and Data Science

Topic 5: Discrete Distributions

William Marsh

Outline

- Aims:
  - Introduce discrete probability distributions

- Coin flipping: Bernoulli distribution
- Multiple flips: the i. i. d. assumption
- Binomial distribution

Flip a Coin

Flip a Coin

- Probability of heads is 50%
  - Provided coin is fair
- Generalise: flip(p)
  - p is probability

$$y_1, y_2, y_3, y_{41} \sim bernoulli(0.5)$$

```
p = 0.5
rv = stats.bernoulli(p)
ys = rv.rvs(4)
```

Aside: random numbers from computers

Repeatedly Flipping

- Repeatedly (10 times) try 4 flips

```
p = 0.5
rv = stats.bernoulli(p)
for x in range(0, 10):
    print(rv.rvs(4))
```

```
[0 0 1 0]
[1 0 1 1]
[0 1 0 1]
[0 0 1 0]
[0 1 1 0]
[0 1 0 0]
[0 0 0 1]
[1 1 1 1]
[1 1 0 0]
[0 1 1 1]
```

Repeatedly Flipping (p=0.3)

- Repeatedly (10 times) try 4 flips

```
p = 0.3
rv = stats.bernoulli(p)
for x in range(0, 10):
    print(rv.rvs(4))
```

```
[0 0 0 0]
[0 0 0 0]
[0 1 0 1]
[0 0 0 1]
[1 1 1 0]
[1 1 1 1]
[0 1 0 0]
[1 1 1 0]
[0 0 0 0]
[0 1 0 1]
```

Probability of Outcome

- Suppose p is 0.3
  - Result of 4 flips: 1, 0, 0, 1
  - Probability?

$$
\begin{aligned}
Pr(1,0,0,1) &= 0.3 \times 0.7 \times 0.7 \times 0.3 \\
&= 0.3^2 \times 0.7^2
\end{aligned}
$$

- Probability does not depend on order
  - Assumed independence (i.i.d)
  - Only count of ‘1’ outcomes

Probably of Outcome II

- Any p, any number of flips N

$$\boxed{Pr(n \text{ from } N \mid p) = p^n.(1-p)^{N-n}}$$

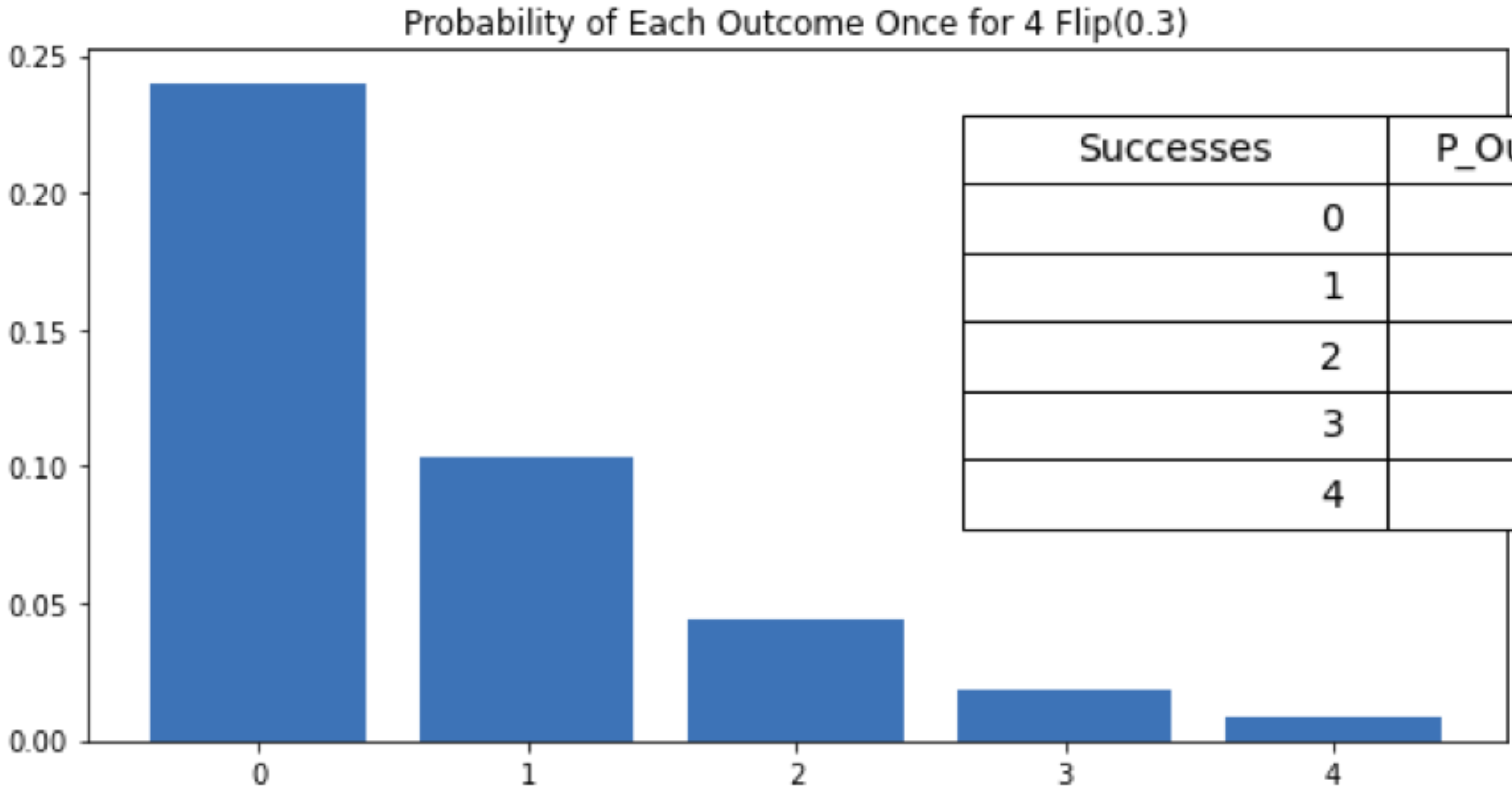
- Sufficient statistic
  - Number of positive ‘n’ outcomes from ‘N’ trials

Binomial Distribution

Probability distribution for outcomes of coin flipping (Bernoulli) trials

Probabilities of Outcomes of 4 Flips

- Calculate probabilities of all outcomes

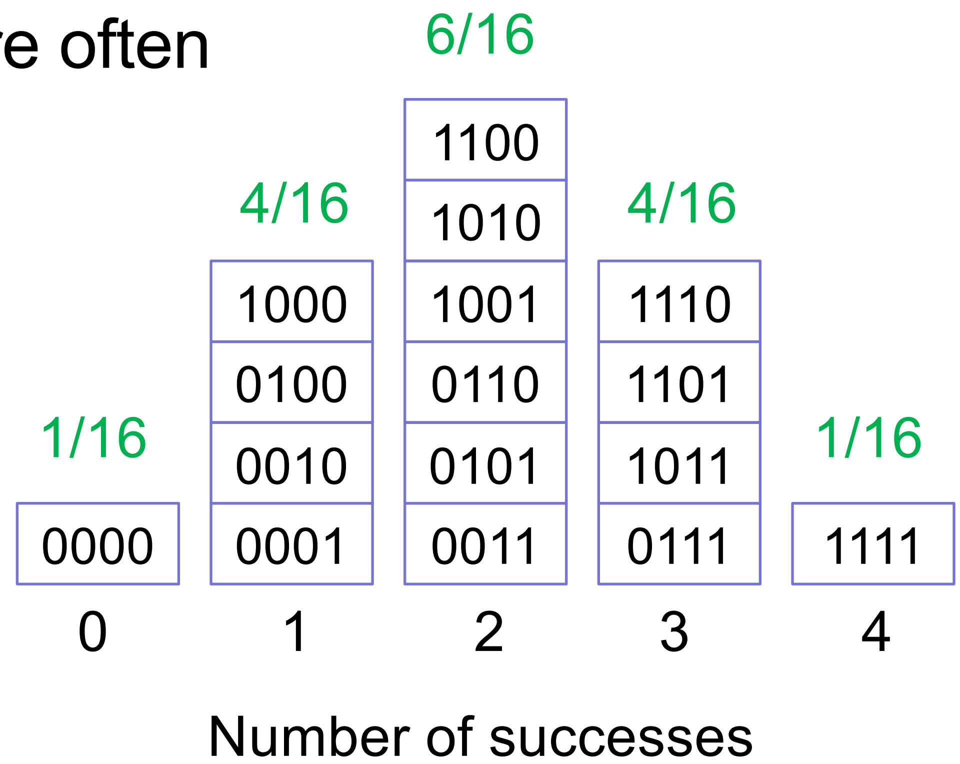


Successes	P_Outcome_Once
0	0.2401
1	0.1029
2	0.0441
3	0.0189
4	0.0081

- But not a probability distribution

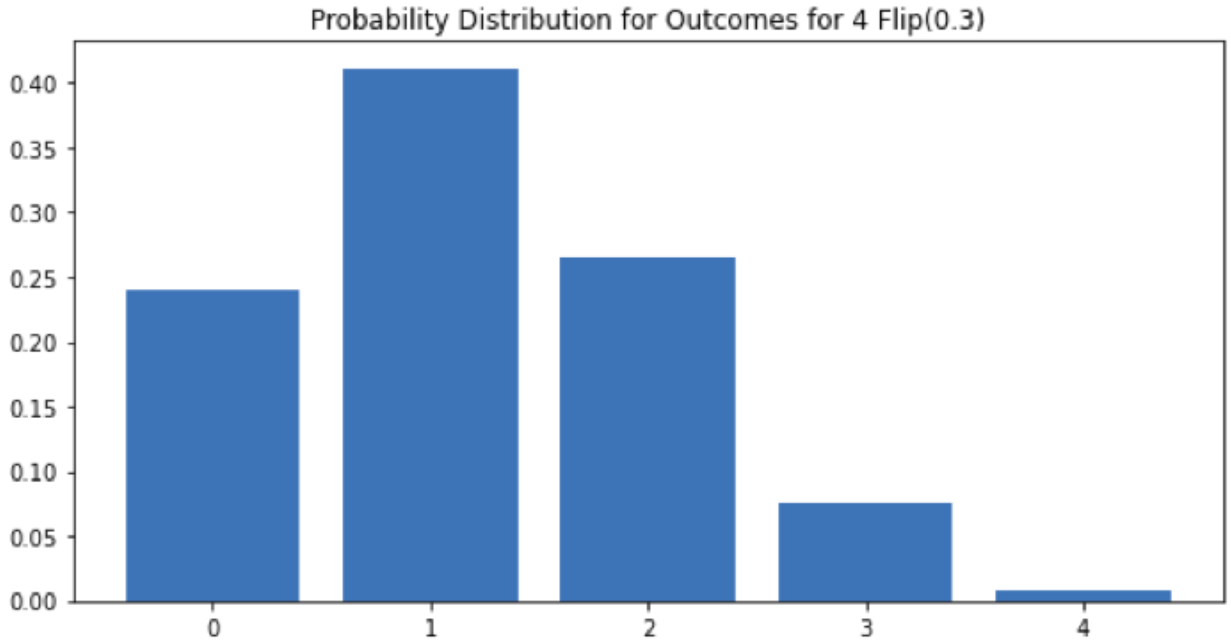
A Trial of 4 Flips

- 16 possible outcomes (incl order)
- 5 different ‘number of successes’
  - Two extremes outcomes (0 or 4 successes) occur once
  - Two success occur more often
- Intuition: 2 heads and 2 tails of a coin more likely



Probability Distribution of Outcomes

- Now a probability distribution
- 'Binomial'



n successes

Pr of one outcome with n successes

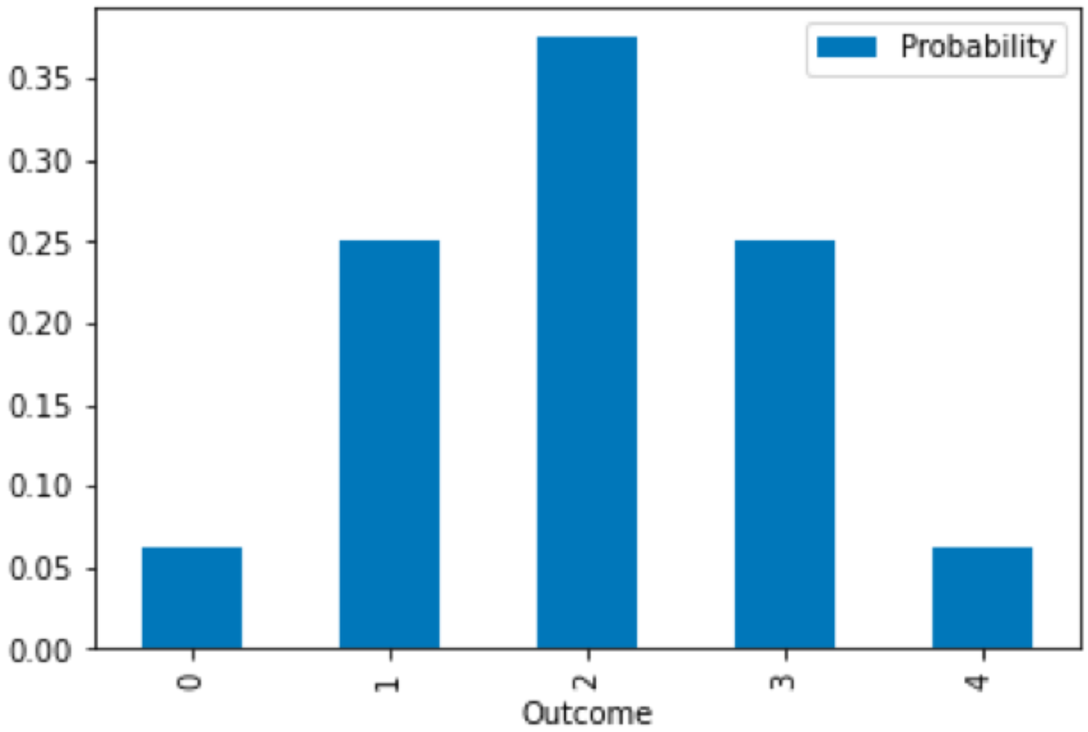
Number of outcomes with n successes

Successes	P_Outcome_Once	Combinations	P_Outcome
0	0.2401	1.0	0.2401
1	0.1029	4.0	0.4116
2	0.0441	6.0	0.2646
3	0.0189	4.0	0.0756
4	0.0081	1.0	0.0081

Outcome probability

Binomial Distribution

- Distribution of outcome of a binary variable
  - Number of ‘trials’ (e.g. coin flips) – N
  - Probability of success (e.g. 50%) – p



binom(4, 0.5)

Sampling

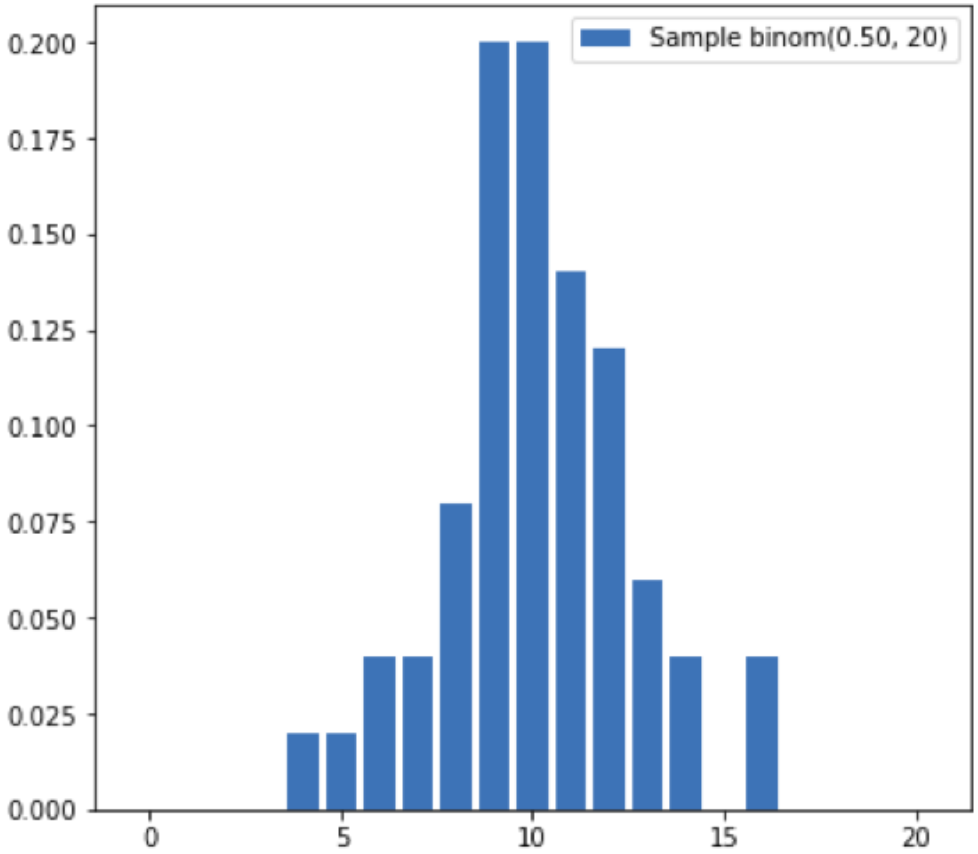
- Flip a coin – sample from Bernoulli(0.5)

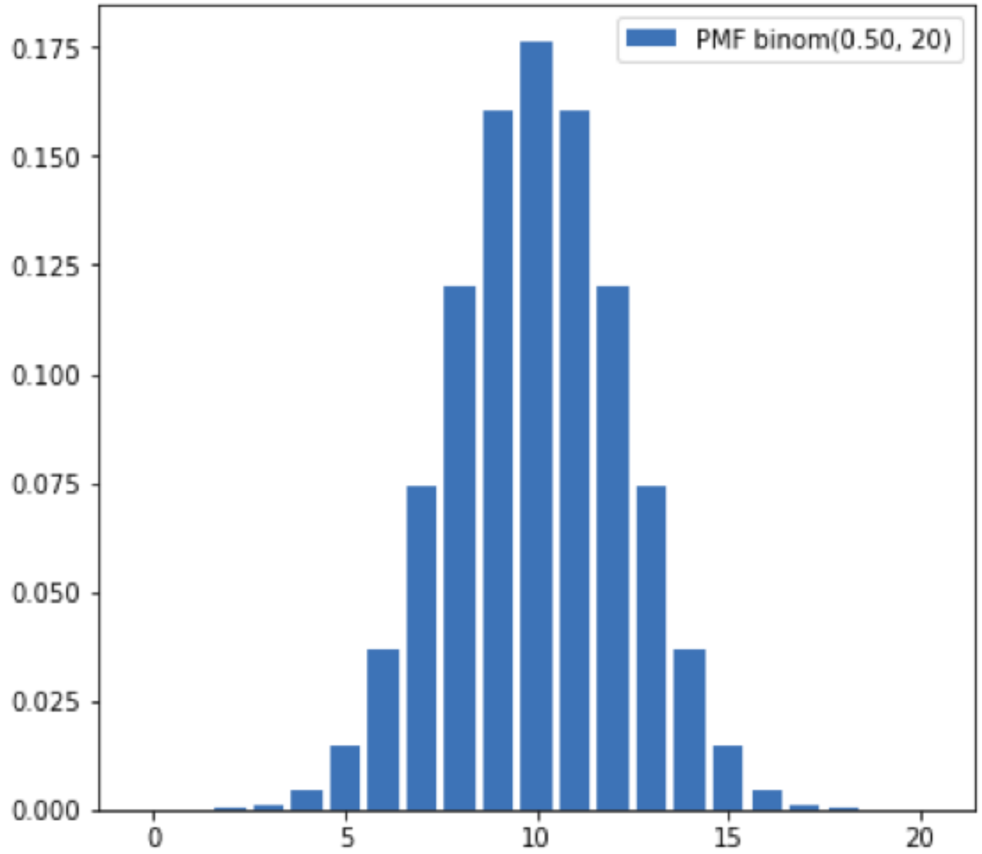
$$coin \sim Bernoulli(0.5)$$

- Run a Bernoulli trial, 5 flips – sample from Binomial

$$successes \sim Binomial(5, 0.5)$$

50 Samples from Binom(20, 0.5)





Quiz 1

Every lecture will have a ‘learning reflection’ slide

Understanding Selection

`.loc[]`

Selection

- The general form of the .loc call is:
  - `.loc[row selector, column selector]`

Use : for all rows

The column selector is ignored if it is omitted

- The 'selectors' can be either:
  - a value or list of values
  - an expression that is true or false

Index!!

expression such as
`df.Area == 'Tower Hamlets’`
Evaluates to a series of True / False

Binomial Probability Mass Function

Binomial PMF

- For a trial of 4 flips, there are 5 possible outcomes
  - 0, 1, 2, 3, 4 successes
  - Distribution shows the probability of each outcome
- Probability of n successes in a trial of N flips, with success probability p:

Num of trial outcomes with n successes

$$\binom{N}{n} p^n (1-p)^{(N-n)}$$

Prob of n successes

Prob of N-n failures

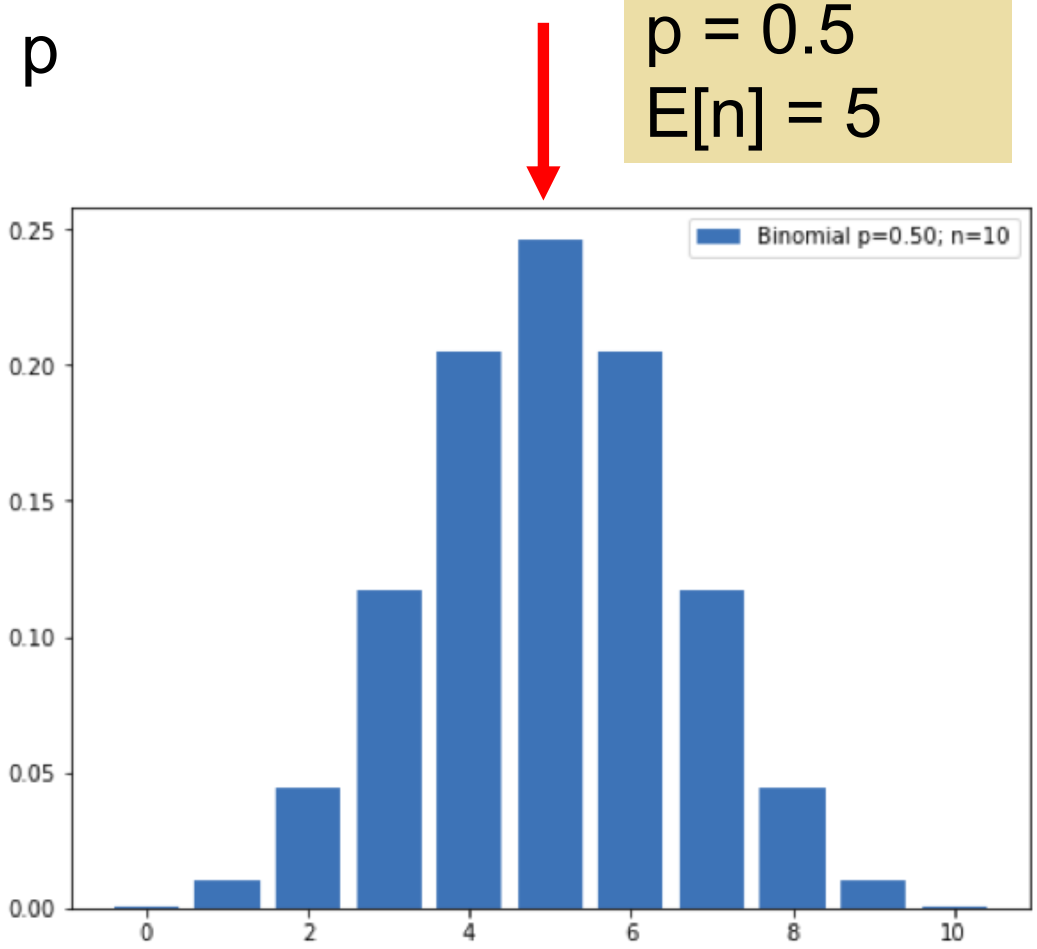
```
N= 10
p = 0.5
stats.binom(N, p).pmf(3)
```

Expected Value and Skew

Expected Value

- Simple formula for expected (average) number of successes
  - Trial of N flips
  - Probability of success p
  - E[n] = N.p

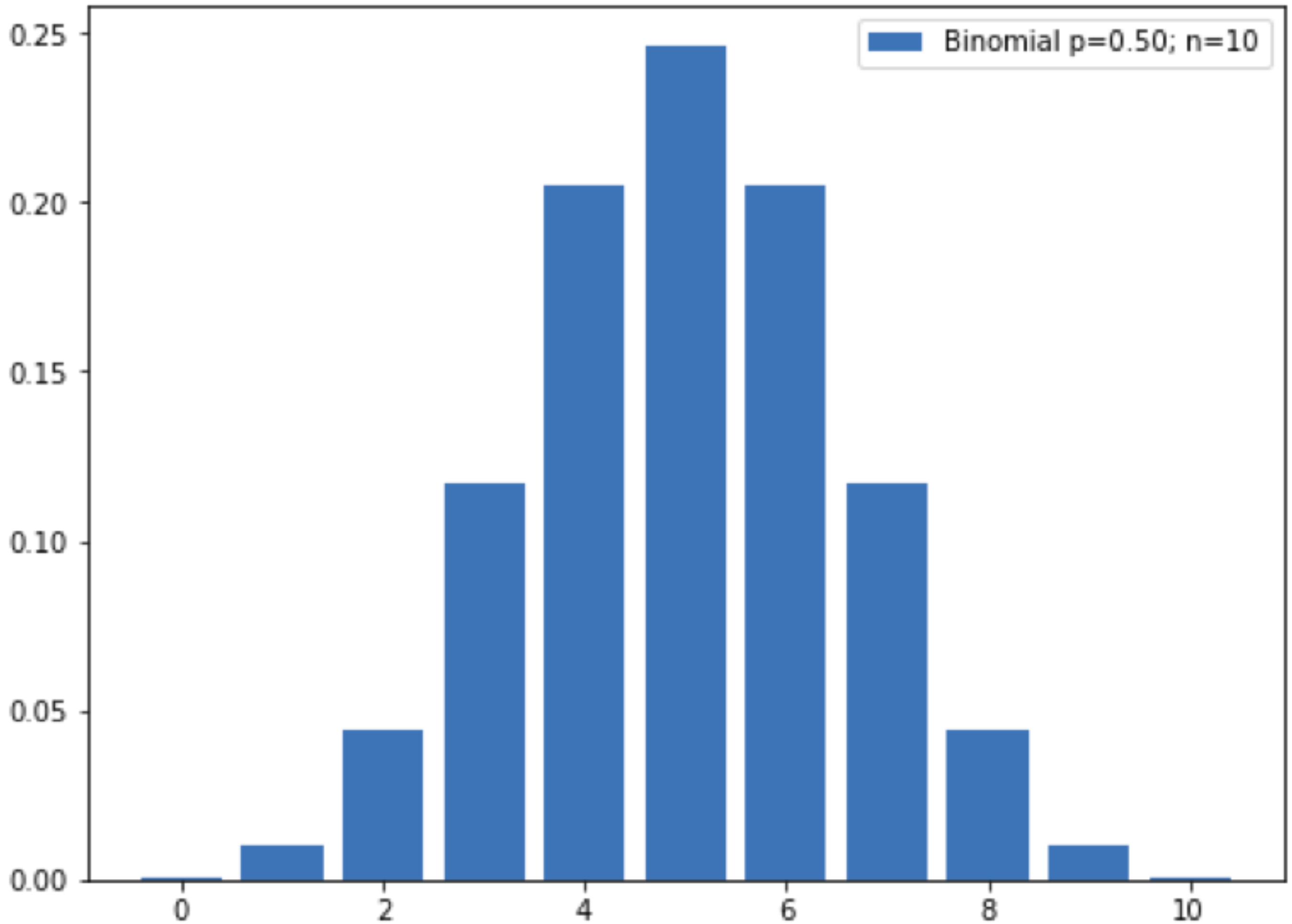
N = 10
p = 0.5
E[n] = 5



Skew of Binomial Distribution

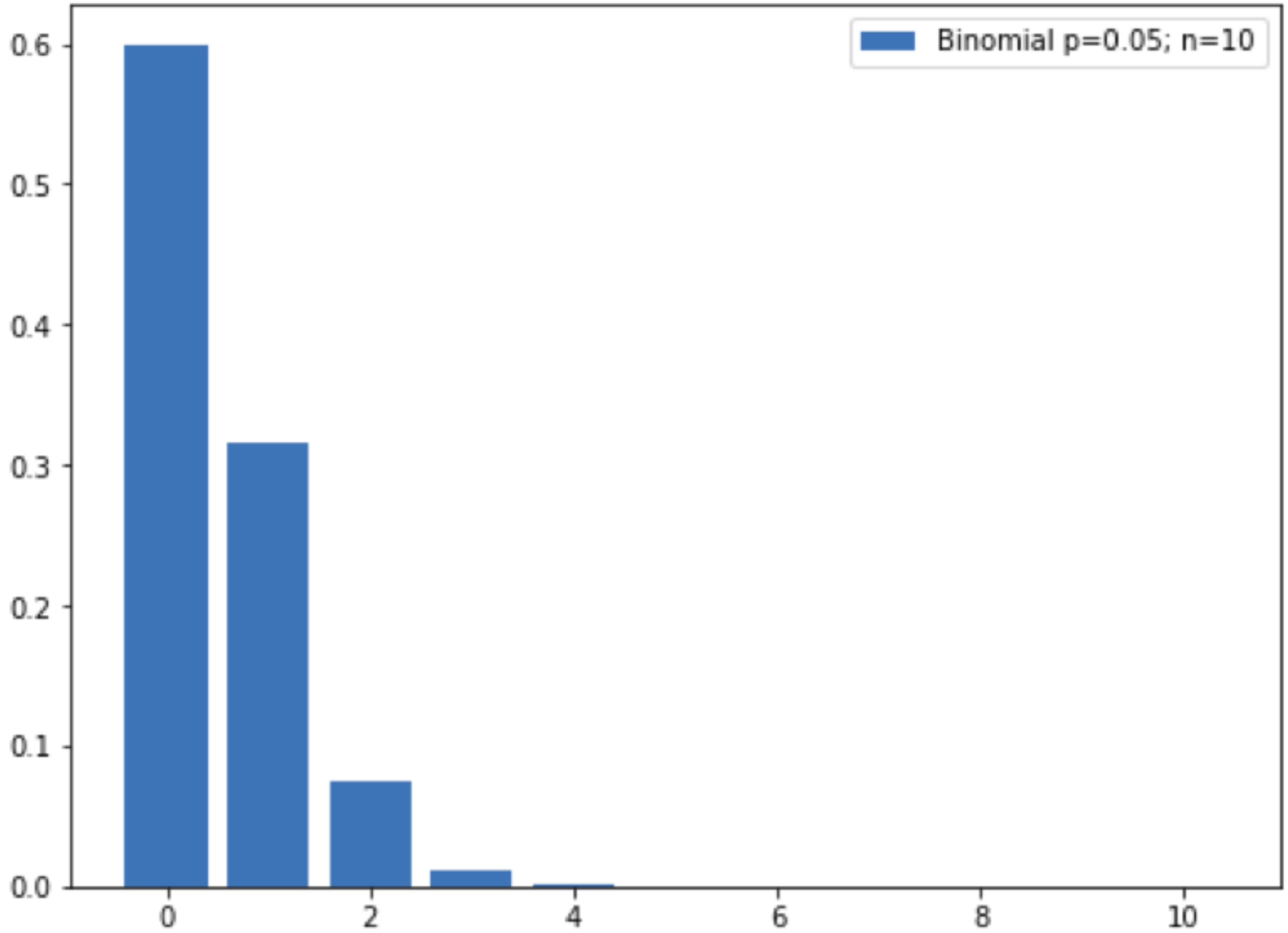
- Skew depends on n and p
- Case 1: expected value close to N/2

N = 10
p = 0.5
E[n] = 5



Skew of Binomial Distribution

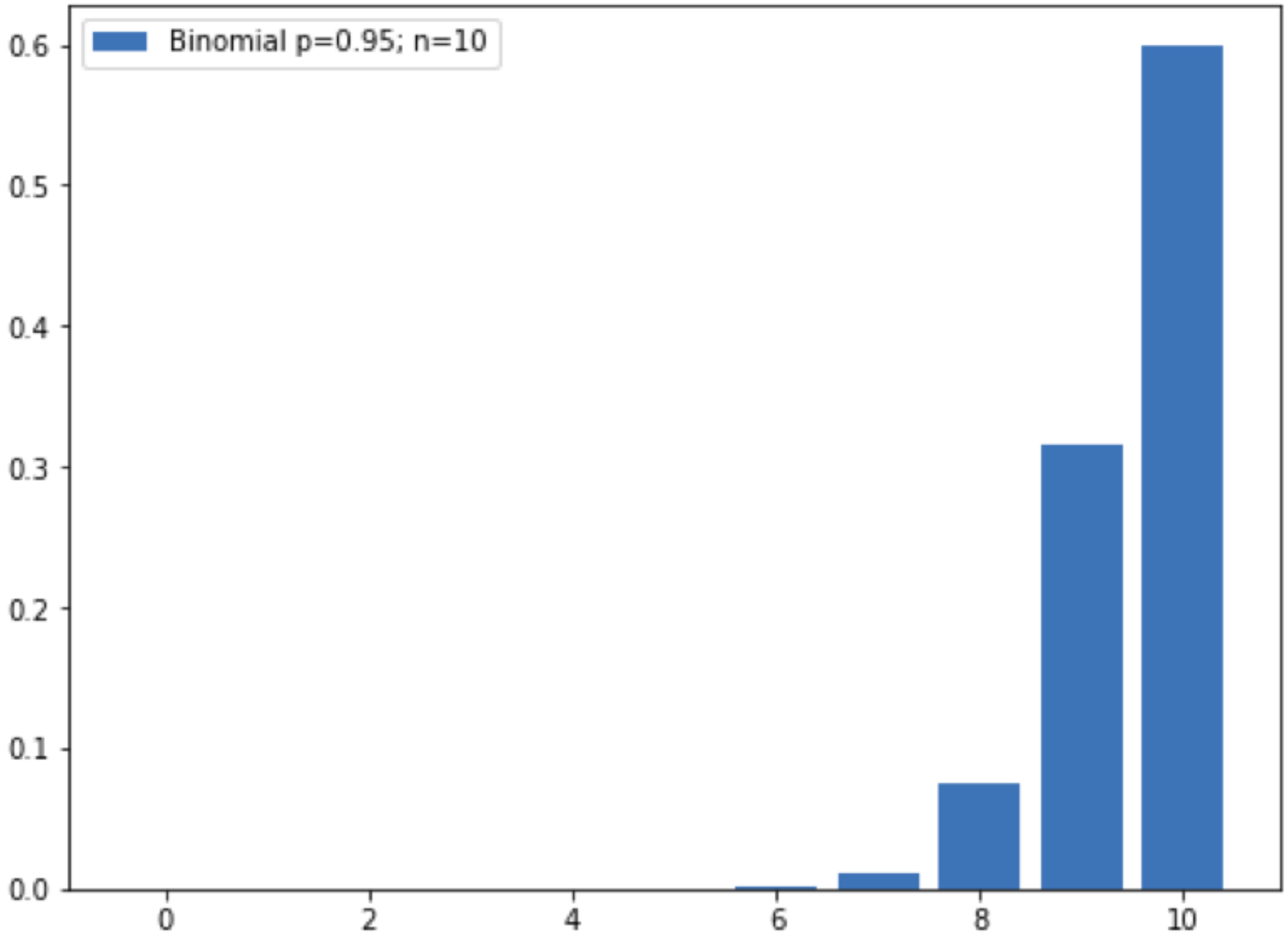
- Skew depends on n and p
- Case 2: expected value close to 0



N = 10
p = 0.05
E[n] = 0.5

Skew of Binomial Distribution

- Skew depends on n and p
- Case 3: expected value close to N



N = 10
p = 0.95
E[n] = 9.5

Poisson Distribution

What if N is not limited?

- A hospital serves a large population
- Most people stay well
- We collect data on the average rate (people / day) of admission to hospital
- What is the distribution of daily admissions?

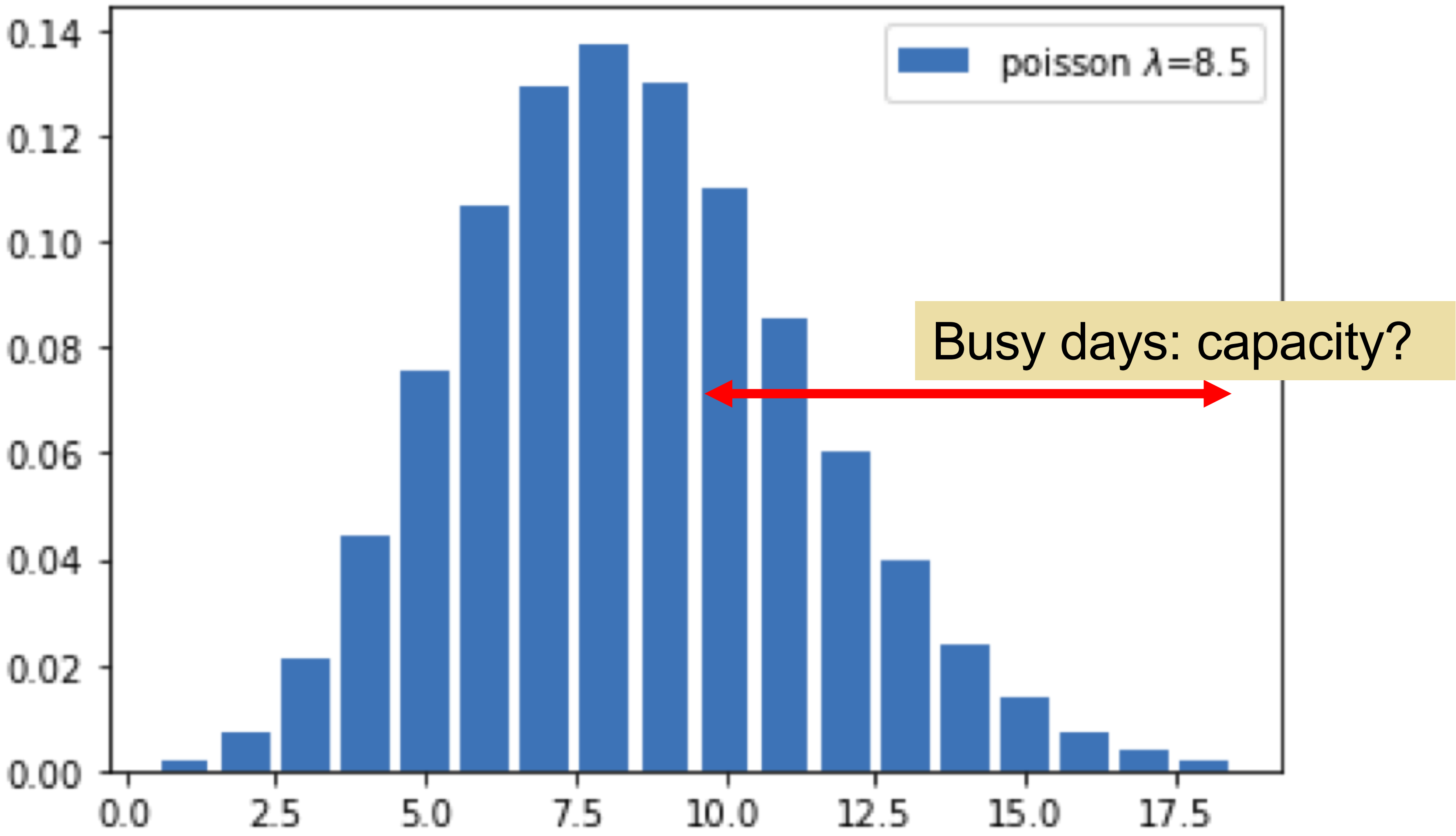
$$x \sim Poisson(\lambda)$$

Integer result: 0 upwards

Rate: real number >= zero

Poisson Example

- Expected value = λ



Quiz 2

Summary

Probability Distribution

- A mathematical formula describing a probability distribution
- Each corresponds to different assumptions about how variation arises

- Some common distributions

Name	Type	Description
categorical(ps)	Discrete	Generalisation of the Bernoulli to >2 outcomes
binomial(n, p)	Discrete	Outcome of n trials, which probability p
poisson(λ)	Discrete	Number of events, occurring at average rate λ
normal(μ, σ)	Continuous	Measurements with mean μ and standard deviation σ
exponential(λ)	Continuous	Time between events, occurring at average rate λ

Summary

- Bernoulli (or flip) distribution
  - Binary outcome – probability p
  - Simplest of all
- Binomial distribution
  - First example of a parametric distribution
  - Result of a Bernoulli trial (flipping repeatedly)
- Idea of sampling