

Conditional probabilities

PREPARING FOR STATISTICS INTERVIEW QUESTIONS IN PYTHON

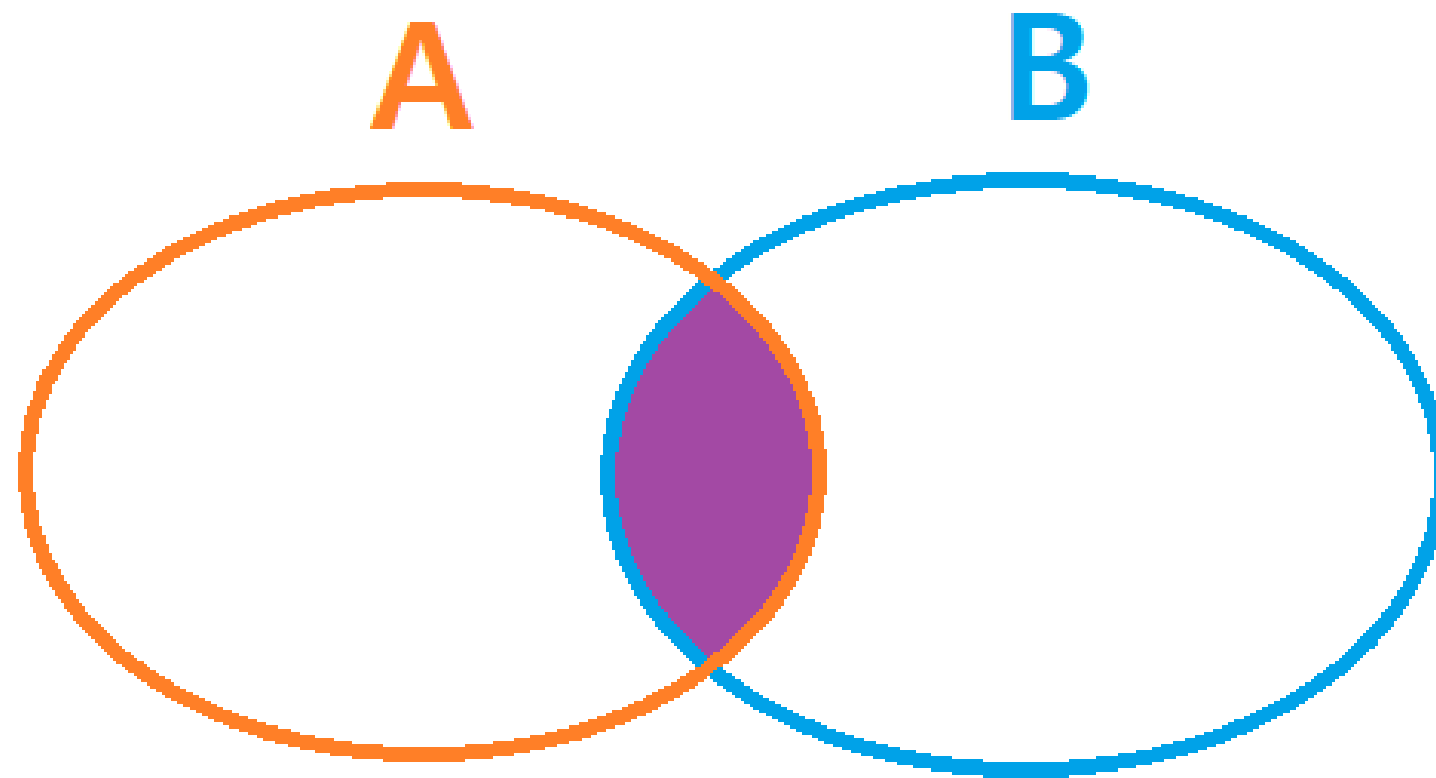


Conor Dewey
Data Scientist, Squarespace

Course overview

1. Probability and sampling distributions
2. Exploratory data analysis
3. Statistical experiments
4. Regression and classification

Quick review



¹ Wikimedia

Bayes' theorem

THE PROBABILITY OF "B"
BEING TRUE GIVEN THAT
"A" IS TRUE

↓

THE PROBABILITY
OF "A" BEING
TRUE

↙

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}$$

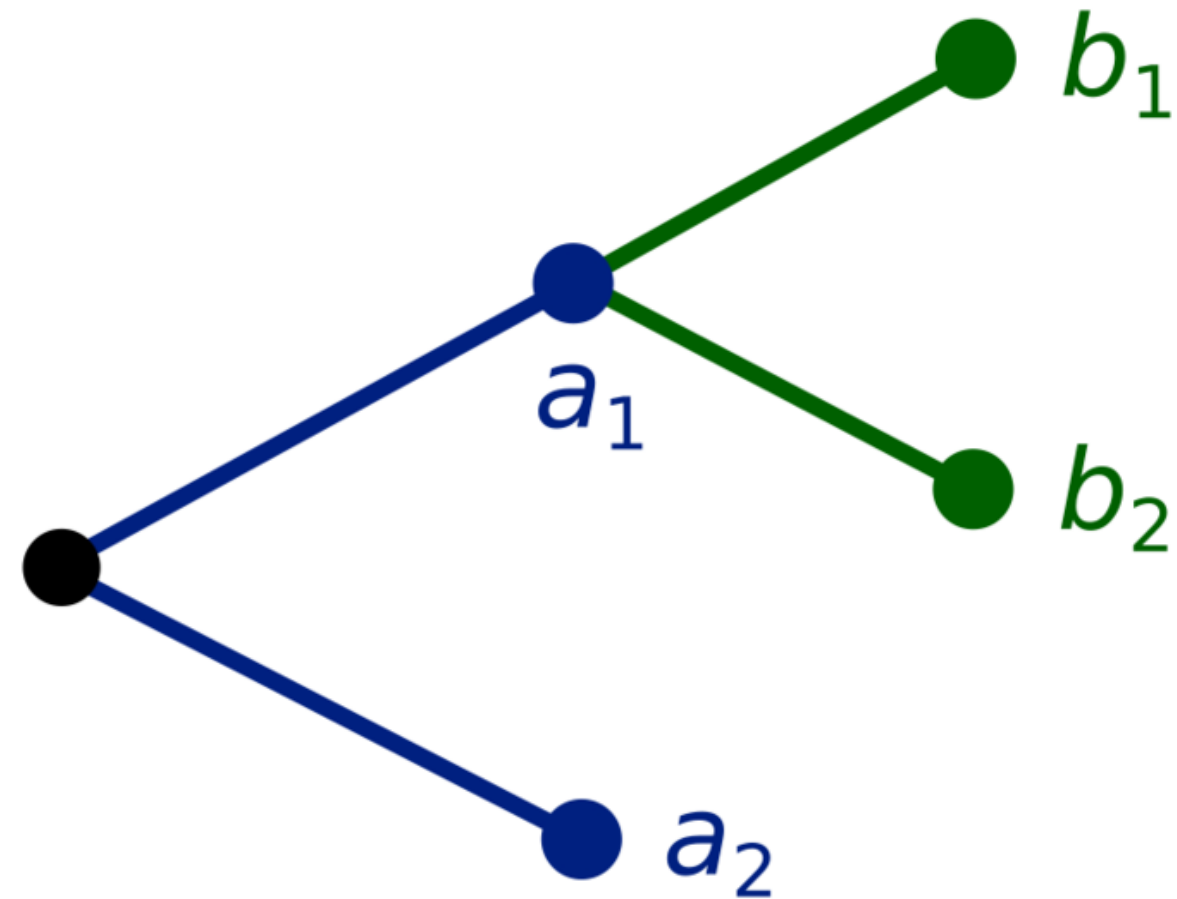
↑

THE PROBABILITY
OF "A" BEING TRUE
GIVEN THAT "B" IS
TRUE

↖

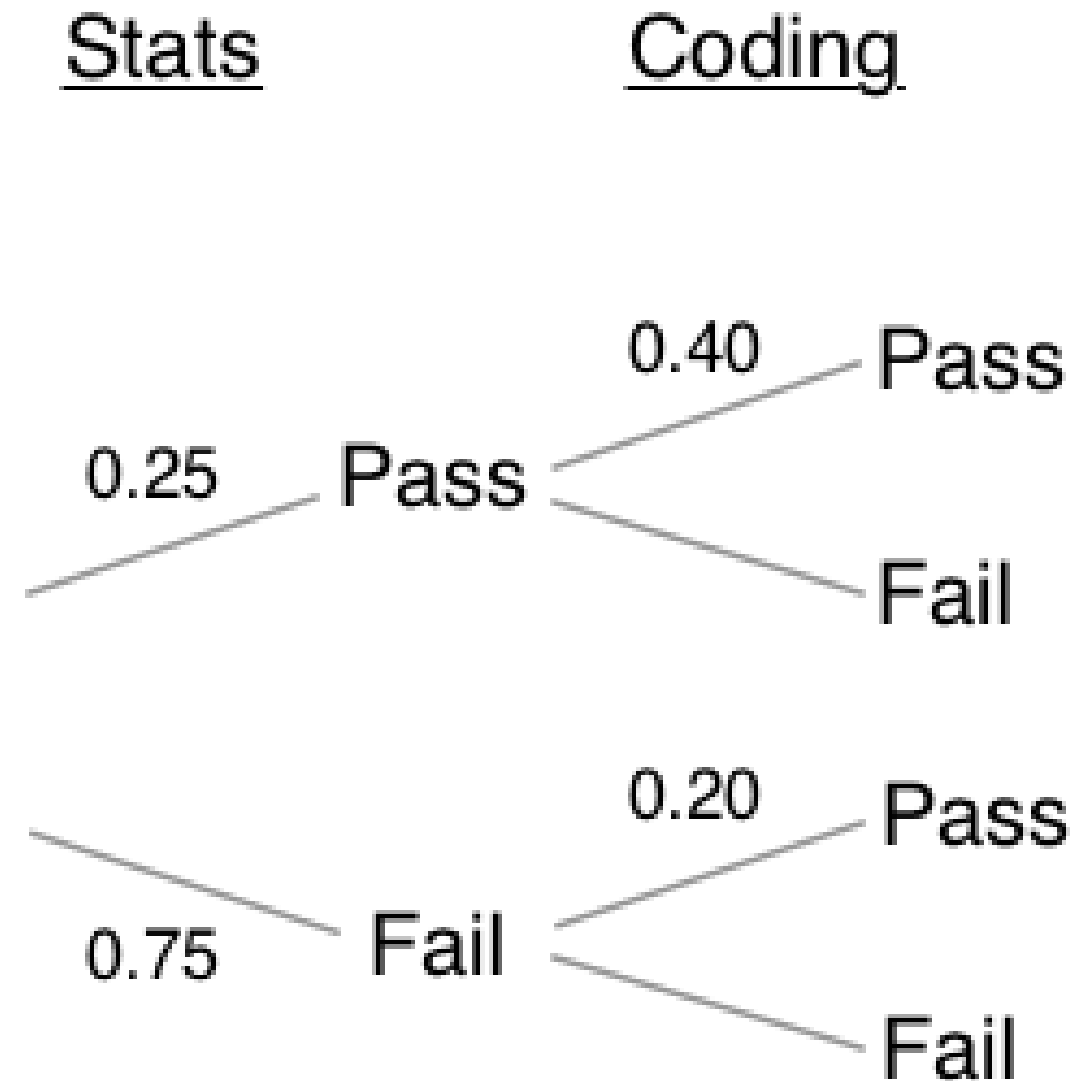
THE PROBABILITY
OF "B" BEING
TRUE

Probability tree diagrams



¹ Wikimedia

Example: passing the interview



Example: passing the interview

```
both = 0.25 * 0.40
```

```
print(both)
```

```
coding = (0.25 * 0.40) + (0.75 * 0.20)
```

```
print(coding)
```

```
0.1
```

```
0.25
```

```
stats_given_coding = both / coding
```

```
print(stats_given_coding)
```

```
0.4
```

Summary

- Conditional probabilities
- Bayes' theorem
- Probability tree diagrams

Let's prepare for the interview!

PREPARING FOR STATISTICS INTERVIEW QUESTIONS IN PYTHON

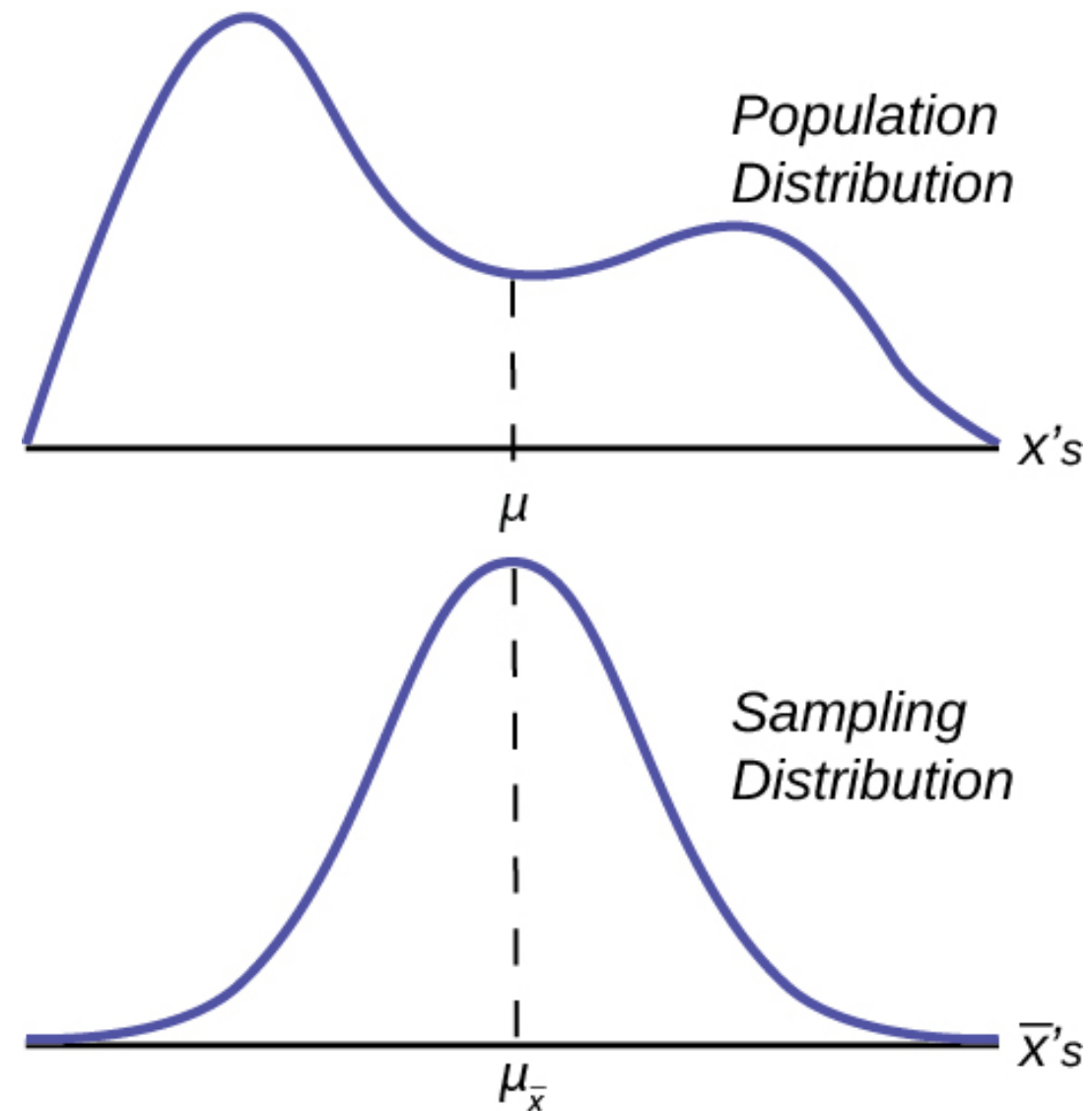
Central limit theorem

PREPARING FOR STATISTICS INTERVIEW QUESTIONS IN PYTHON



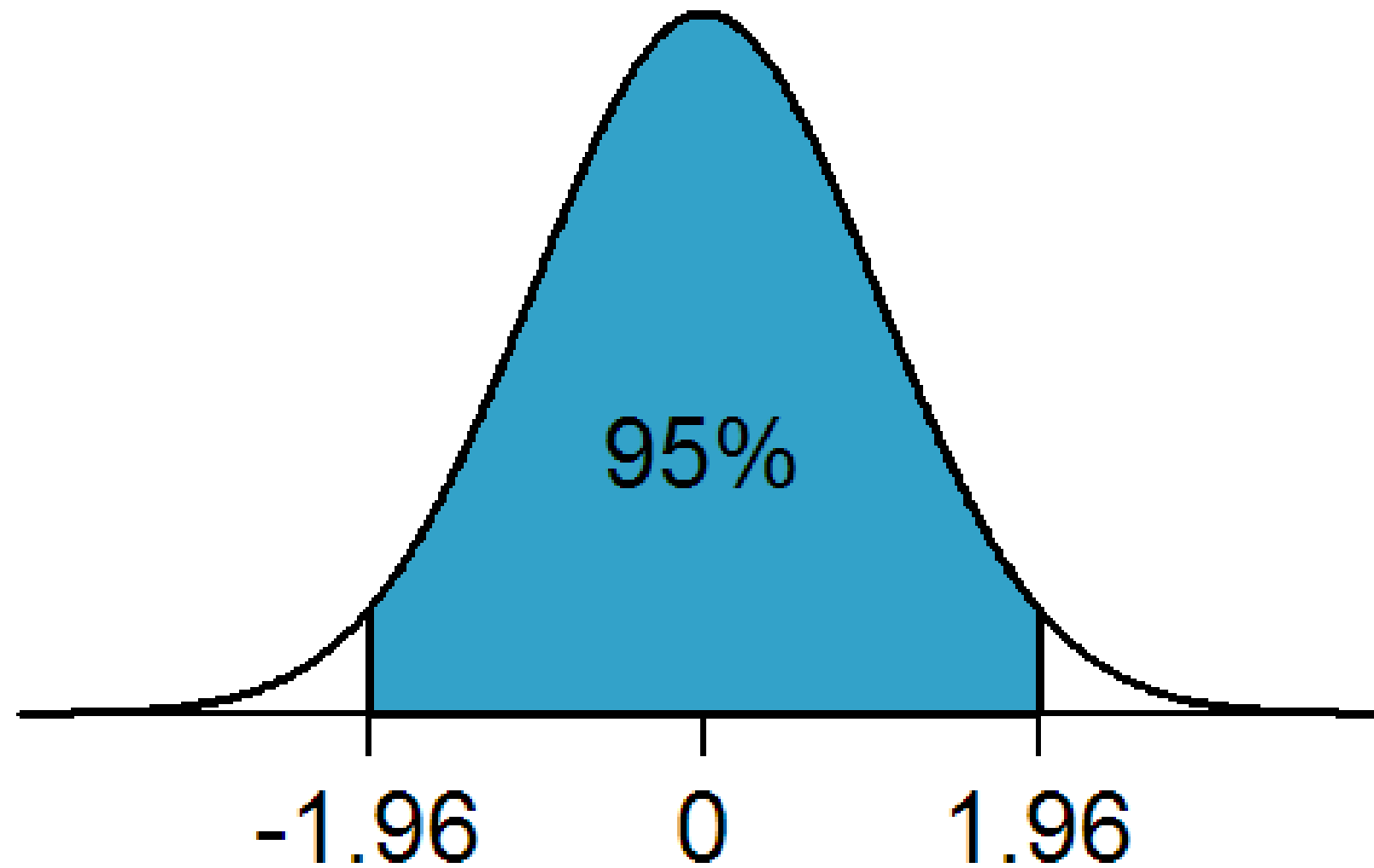
Conor Dewey
Data Scientist, Squarespace

What does it mean?



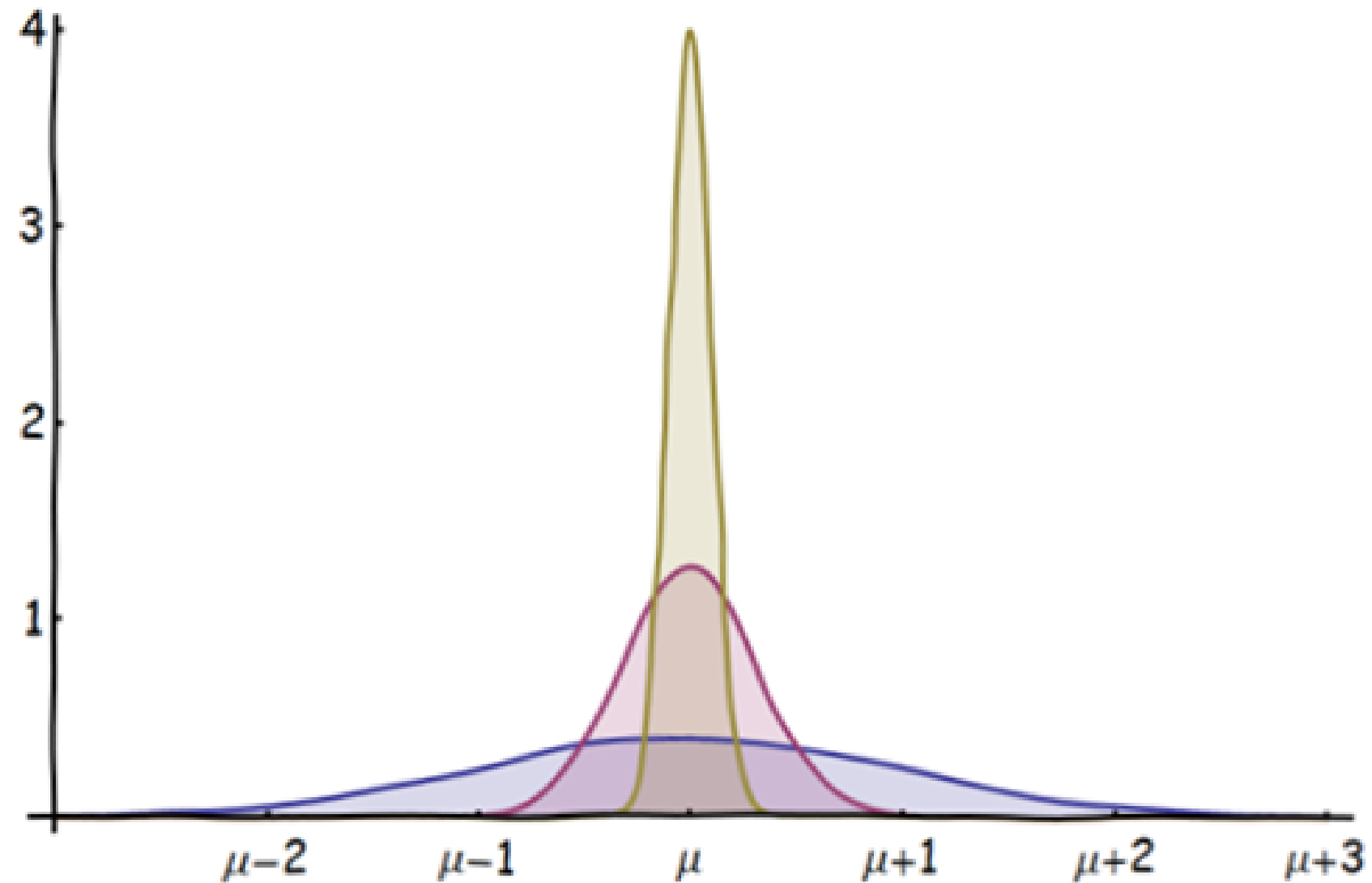
¹ OpenStax CNX

Why does it matter?



¹ Wikimedia

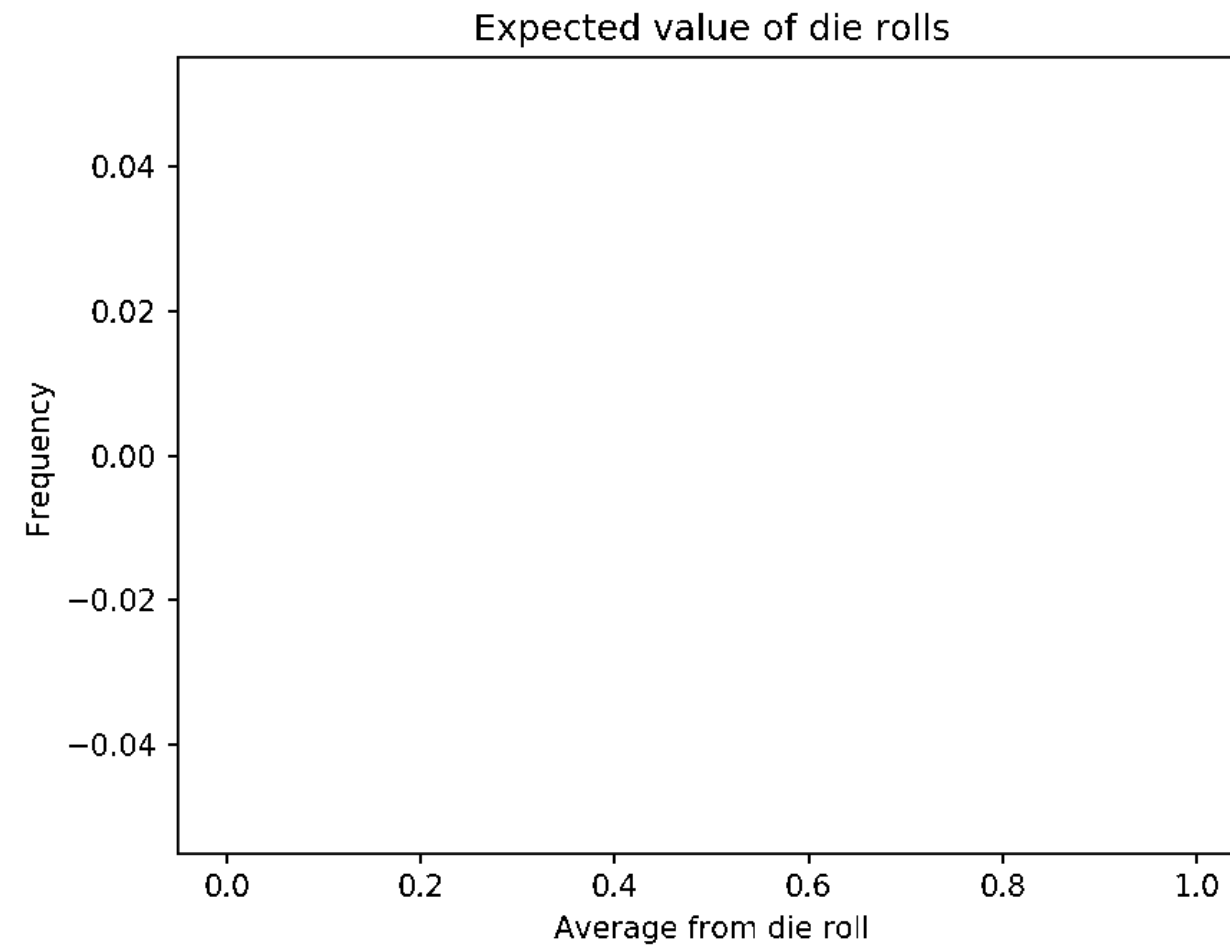
Law of large numbers



¹ StackExchange

Simulating CLT in Python

```
np.random.randint(start, end, size)
```



¹ How to Visualize the Central Limit Theorem in Python

List comprehension

```
x = [1,2,3,4]
out = []
for item in x:
    out.append(item**2)
print(out)
```

```
[1, 4, 9, 16]
```

```
x = [1,2,3,4]
out = [item**2 for item in x]
print(out)
```

```
[1, 4, 9, 16]
```

Summary

- Central limit theorem
- Law of large numbers
- Simulating die rolls
- List comprehension

Let's prepare for the interview!

PREPARING FOR STATISTICS INTERVIEW QUESTIONS IN PYTHON

Probability distributions

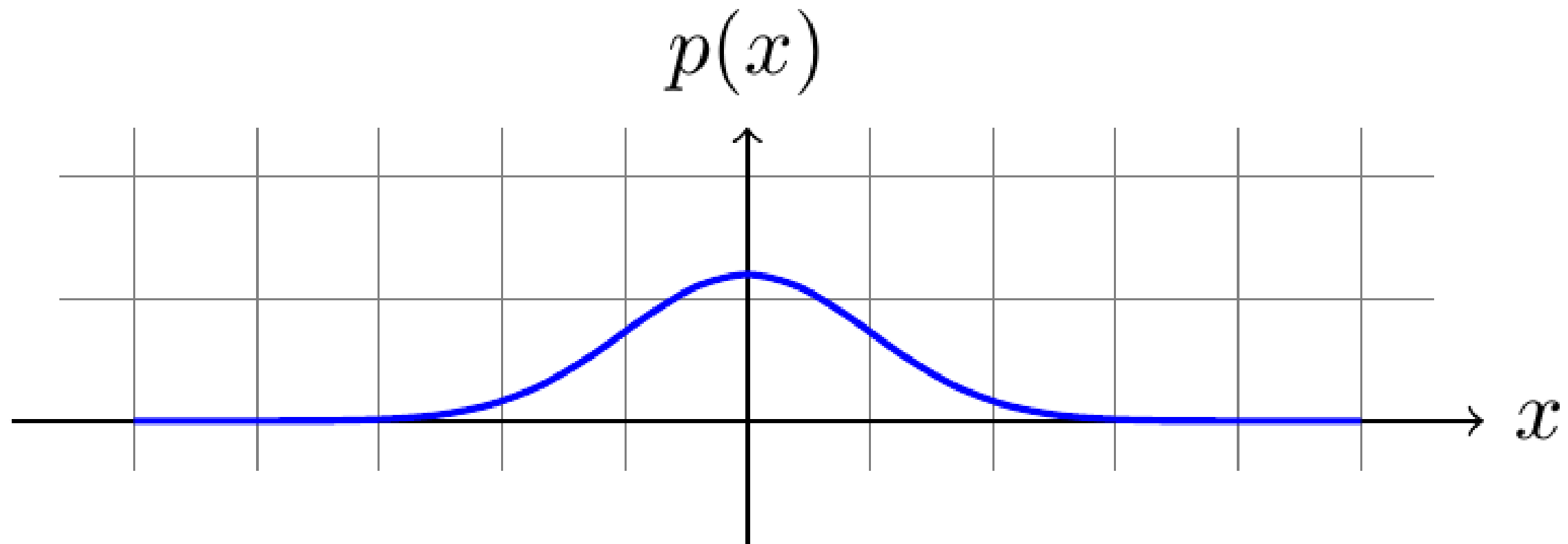
PREPARING FOR STATISTICS INTERVIEW QUESTIONS IN PYTHON



Conor Dewey
Data Scientist, Squarespace

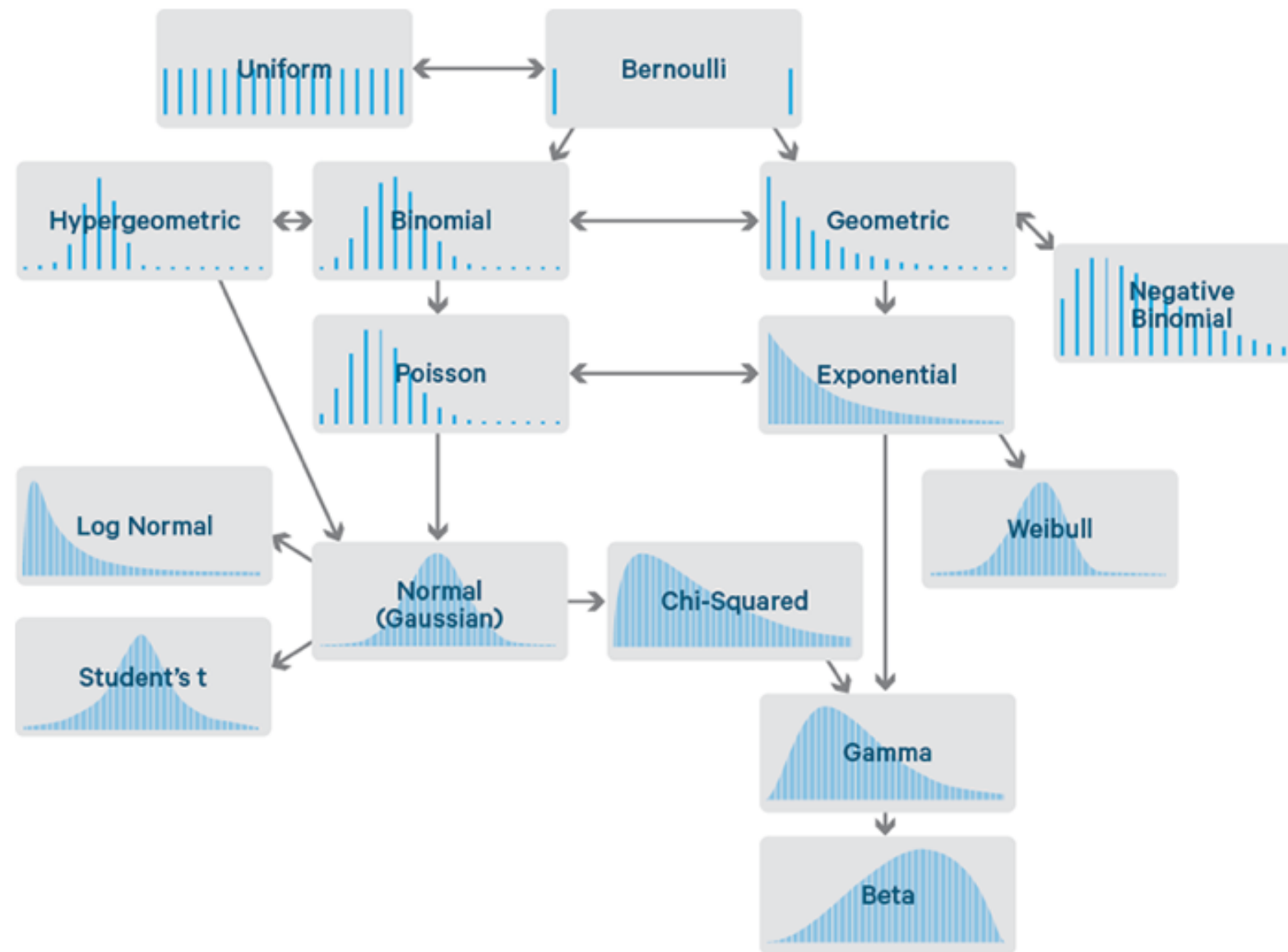
What's a probability distribution?

- Indicates likelihood of an outcome
- Probabilities must add up to 1



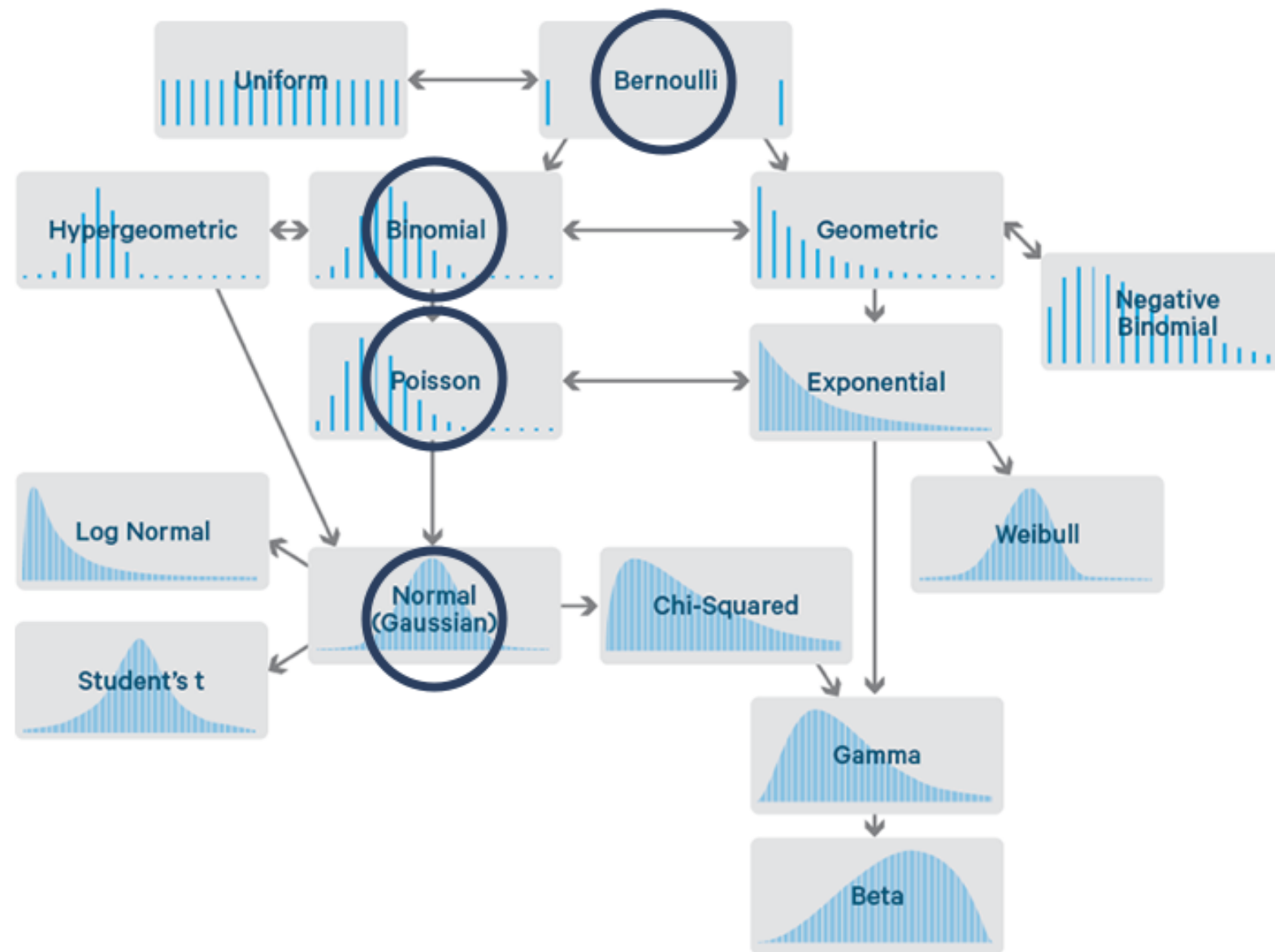
¹ Wikimedia

Overview of common distributions



¹ Common Probability Distributions: The Data Scientists Crib Sheet

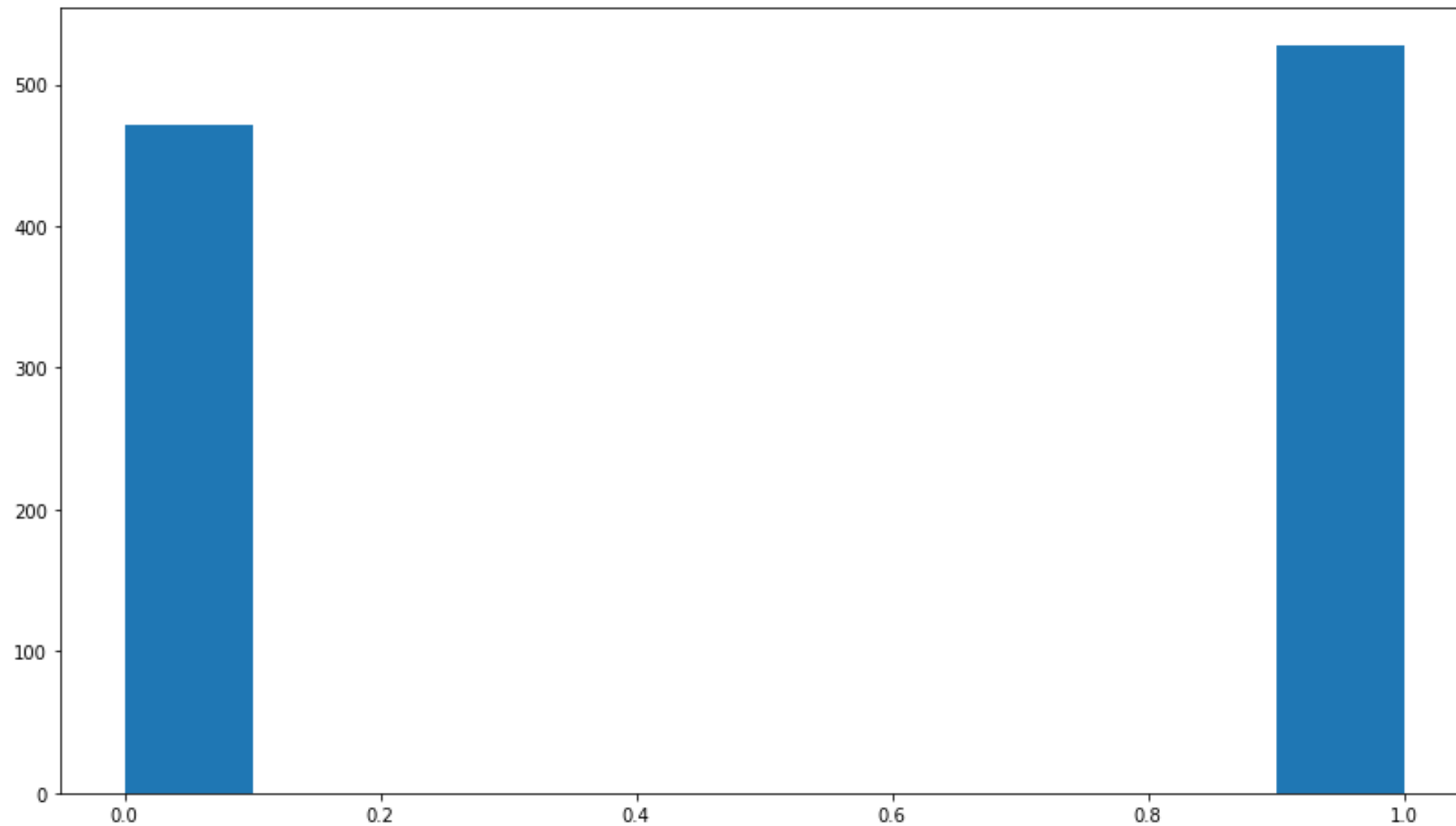
Overview of common distributions



¹ Common Probability Distributions: The Data Scientists Crib Sheet

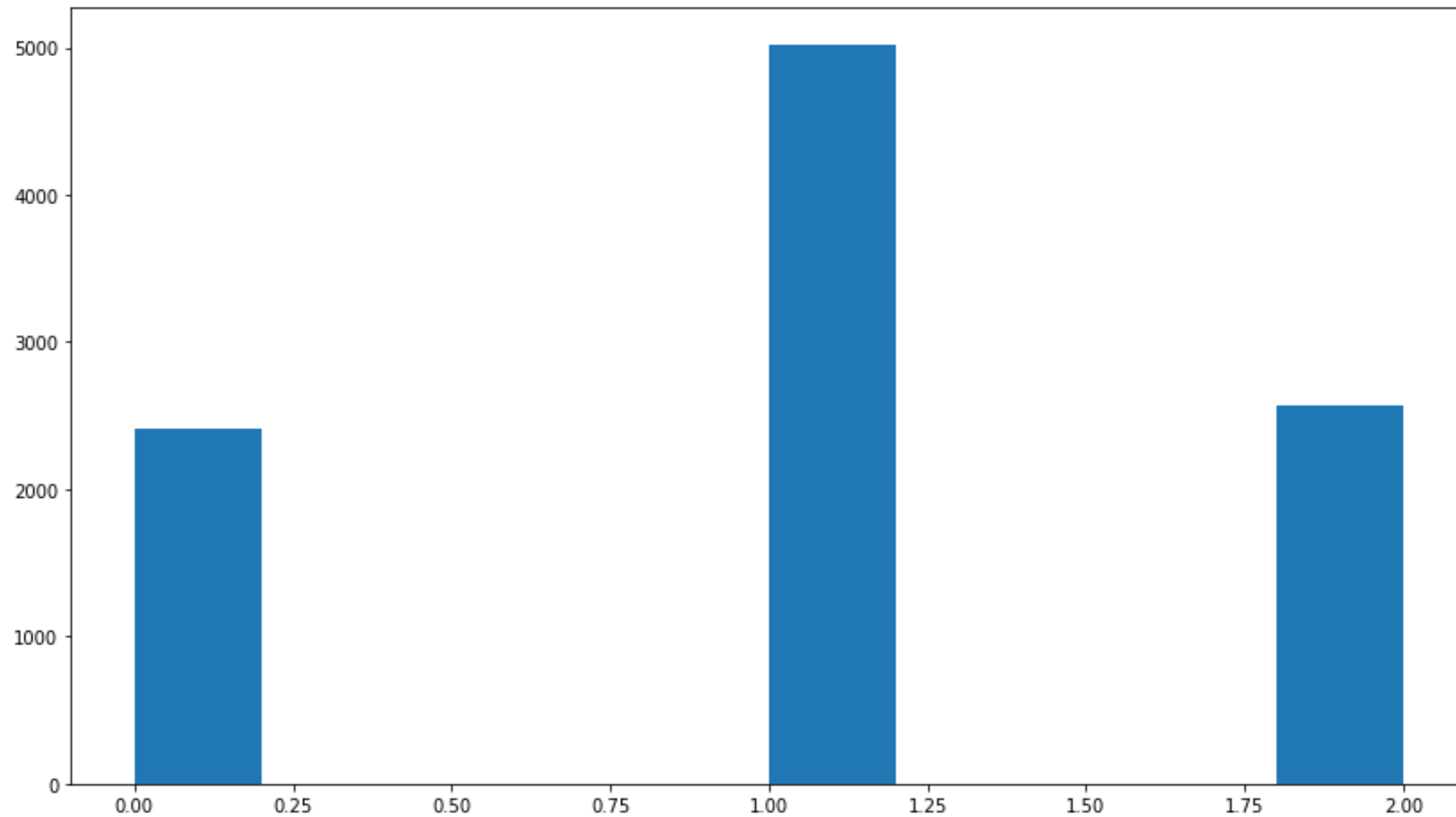
Bernoulli distribution

```
plt.hist(bernoulli.rvs(p=0.5, size=1000))
```

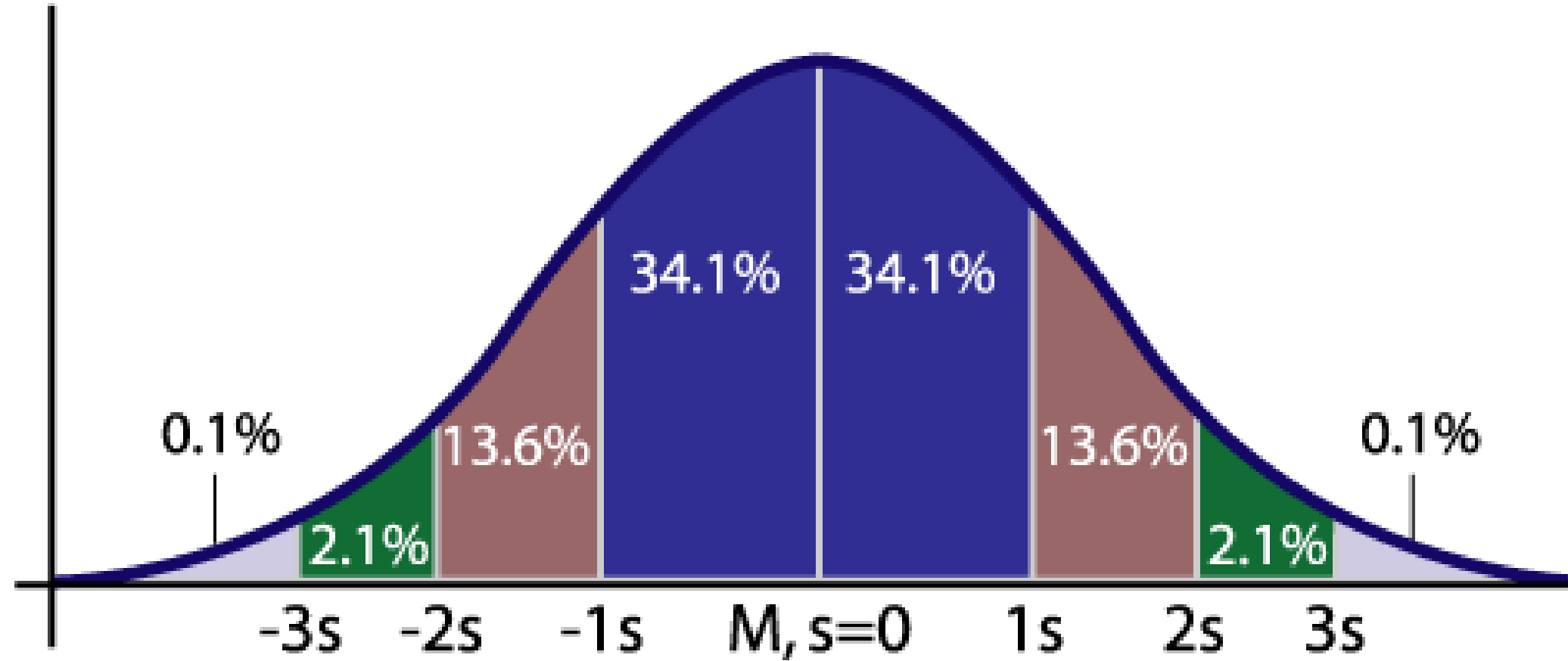


Binomial distribution

```
plt.hist(binom.rvs(2, 0.5, size=10000))
```

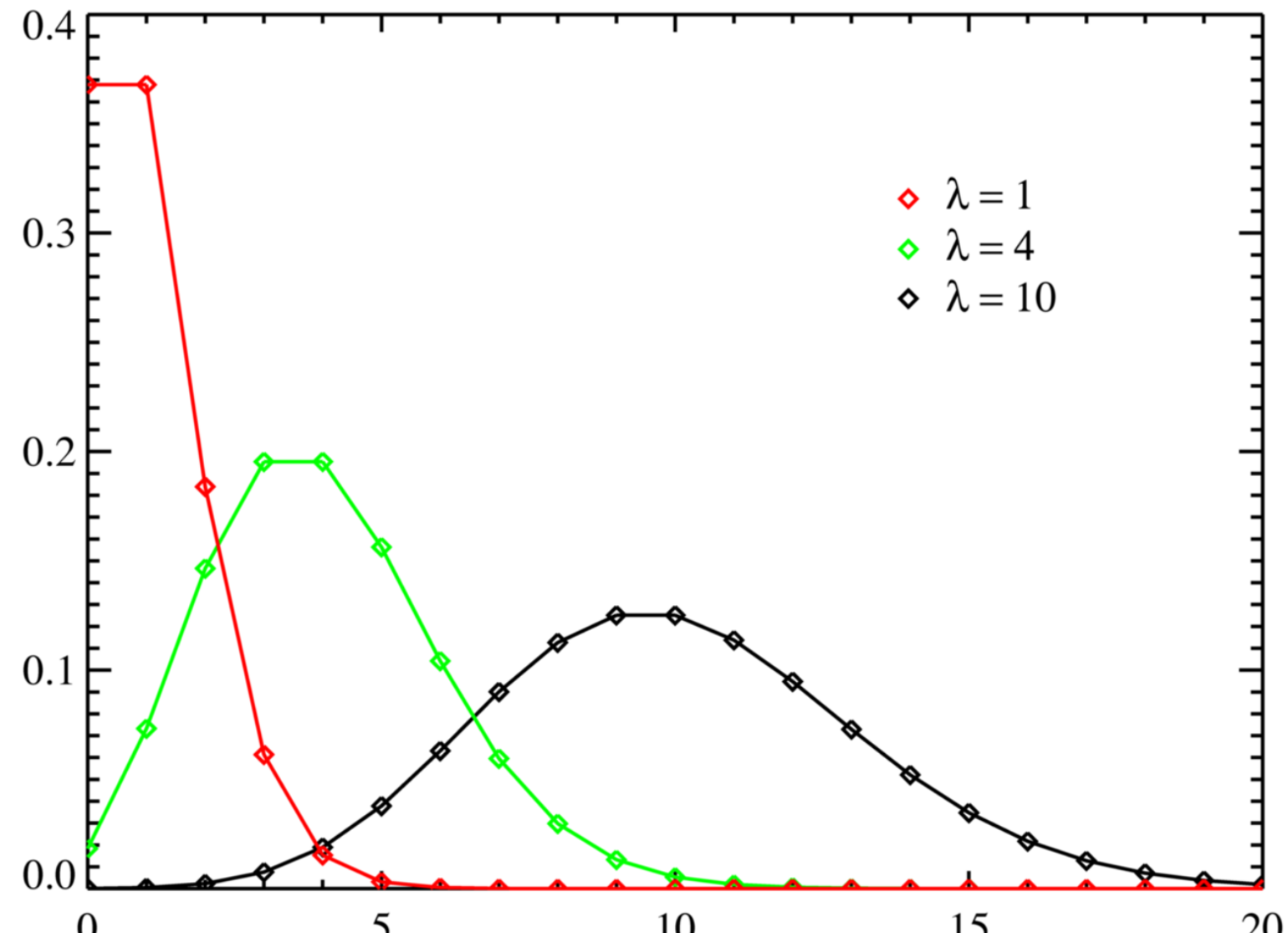


Normal distribution



¹ Wikimedia

Poisson distribution



¹ Wikimedia

Poisson distribution

In any 15-minute interval, there is a 20% probability that you will see at least one shooting star. What is the probability that you see at least one shooting star in the period of an hour?

Summary

- Definition of probability distributions
- Overview of common distributions
- Bernoulli, binomial, normal, and Poisson

Let's prepare for the interview!

PREPARING FOR STATISTICS INTERVIEW QUESTIONS IN PYTHON