

Home

Workspace

Recents

Data

Clusters

Jobs

Search

3.2.1 Lab - Bernoulli Distribution (Python)

Detached

File

Edit

View: Standard

Permissions

Run All

Clear

Publish

Comments

Experiment

Revision history

Cmd 1

Cmd 2

Bernoulli Distribution (20 mins)

Objective: Use a Bernoulli distribution to model a real-world scenario.

In this lab, you will complete a series of exercises where you model a real-world scenario using the Bernoulli distribution.

Cmd 3

Exercise 1

Random variable X represents whether we roll a number divisible by 3 on a six sided die.

- Let x = 1 represent success (rolling a 3 or 6)
- Let x = 0 represent failure (rolling a 1, 2, 4, or 5)

The probability of an outcome n can be computed using the probability density function of a Bernoulli distribution:

$P(x) = p^x(1-p)^{1-x}$, where p is the probability of success.

Compute the probability of success (x=1) using this function.

Cmd 4

```
1 # ANSWER
2 p = 1/3
3 x = 1
4 prob_success = p**x * (1-p)**(1-x)
5 print(f"The probability of success is {prob_success}")
```

Cmd 5

Exercise 2

Use the same function to compute the probability of failure (x=0).

Cmd 6

```
1 # ANSWER
2 p = 1/3
3 x = 0
4 prob_failure = p**x * (1-p)**(1-x)
5 print(f"The probability of failure is {prob_failure}")
```

Cmd 7

Exercise 3

Get the expected value and variance of this distribution.

- expected value = p
- variance = p(1-p)

Cmd 8

```
1 # ANSWER
2 p = 1/3
3 expected_value = p
4 variance = p * (1-p)
5
6 print(f"The expected value is {expected_value} and the variance is {variance}")
```

Cmd 9

Exercise 4

So far, we've been working with probabilities using Python as a calculator. However, we can also work with probabilities using the `scipy.stats` module.

Complete the below code block to compute the expected value (mean) and variance using `scipy.stats`.

Hint: You can use the [documentation](#) as a reference.

Cmd 10

```
1 # ANSWER
2 from scipy.stats import bernoulli
3 expected_value, variance = bernoulli.stats(p, moments="mv")
4 print(f"The expected value is {expected_value} and the variance is {variance}")
```

Cmd 11

Congratulations on completing the Bernoulli Distribution lab!

Next, we will continue by learning about continuous probability. Learning to use Python modules like `scipy.stats` will be incredibly helpful when working with continuous probability.

Cmd 12

© 2021 Databricks, Inc. All rights reserved.

Apache, Apache Spark, Spark and the Spark logo are trademarks of the [Apache Software Foundation](#).

[Privacy Policy](#) | [Terms of Use](#) | [Support](#)

Shift+Enter to run