

How can conditional probability be calculated in Python?

Authored by
stats writer

May 6, 2024

RECOMMENDED CITATION

stats writer (2024). *How can conditional probability be calculated in Python?*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=143248>

Conditional probability is a mathematical concept that calculates the probability of an event occurring given that another event has already happened. In Python, this can be calculated using the conditional probability formula or by using built-in functions such as "if" statements and "and" operators. The steps involved in calculating conditional probability in Python include defining the events, determining the sample space, and using the appropriate formula or function. By incorporating conditional probability calculations in Python, users can analyze and predict the likelihood of events occurring in a given scenario, making it a valuable tool in data analysis and decision making.

Calculate Conditional Probability in Python

The conditional probability that event A occurs, given that event B has occurred, is calculated as follows:

$$P(A|B) = P(A \cap B) / P(B)$$

where:

$P(A \cap B)$ = the probability that event A and event B both occur.

$P(B)$ = the probability that event B occurs.

The following example shows how to use this formula to calculate conditional probabilities in Python.

Example: Calculate Conditional Probability in Python

Suppose we send out a survey to 300 individuals asking them which sport they like best: baseball, basketball,

football, or soccer.

We can create the following table in Python to hold the survey responses:

```
import pandas as pd
import numpy as np

#create pandas DataFrame with raw data
df = pd.DataFrame({'gender': np.repeat(np.array(), 150),
'sport': np.repeat(np.array(),
(34, 40, 58, 18, 34, 52, 20, 44))})

#produce contingency table to summarize raw data
survey_data = pd.crosstab(index=df, columns=df,
margins=True)

#view contingency table
survey_data

sport Baseball Basketball Football Soccer All
gender
Female 34 52 20 44 150
Male 34 40 58 18 150
All 68 92 78 62 300
```

Related:

We can use the following syntax to extract values from the table:

```
#extract value in second row and first column  
survey_data.iloc
```

34

We can use the following syntax to calculate the probability that an individual is male, given that they prefer baseball as their favorite sport:

```
#calculate probability of being male, given that  
individual prefers baseball  
survey_data.iloc / survey_data.iloc
```

0.5

And we can use the following syntax to calculate the probability that an individual prefers basketball as their favorite sport, given that they're female:

```
#calculate probability of preferring basketball, given
```

that individual is female

survey_data.iloc / survey_data.iloc

0.34666666666666667

The following tutorials provide additional information on dealing with probability:

ARABPSYCHOLOGY.COM