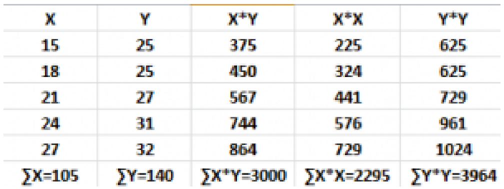
Program to find correlation coefficient

Given two array elements and we have to find the correlation coefficient between two array. Correlation coefficient is an equation that is used to determine the strength of relation between two variables. Correlation coefficient sometimes called as cross correlation coefficient. Correlation coefficient always lies between -1 to +1 where -1 represents X and Y are negatively correlated and +1 represents X and Y are positively correlated.

$$r = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2][n\Sigma y^2 - (\Sigma y)^2]}}$$

Where r is correlation coefficient.



X	Y	X*Y	X*X	Y*Y
15	25	375	225	625
18	25	450	324	625
21	27	567	441	729
24	31	744	576	961
27	32	864	729	1024
∑X=105	∑Y=140	∑X*Y=3000	∑X*X=2295	∑Y*Y=3964

```
Correlation coefficient
= (5 * 3000 - 105 * 140)
     / sqrt((5 * 2295 - 105²)*(5*3964 - 140²))
= 300 / sqrt(450 * 220) = 0.953463
```

Examples :

```
Input : X[] = {43, 21, 25, 42, 57, 59}
        Y[] = {99, 65, 79, 75, 87, 81}
Output : 0.529809

Input : X[] = {15, 18, 21, 24, 27};
        Y[] = {25, 25, 27, 31, 32}
Output : 0.953463
```

C++

```
// Program to find correlation coefficient
#include<bits/stdc++.h>
```

```
using namespace std;

// function that returns correlation coefficient.
float correlationCoefficient(int X[], int Y[], int n)
{

    int sum_X = 0, sum_Y = 0, sum_XY = 0;
    int squareSum_X = 0, squareSum_Y = 0;

    for (int i = 0; i < n; i++)
    {
        // sum of elements of array X.
        sum_X = sum_X + X[i];

        // sum of elements of array Y.
        sum_Y = sum_Y + Y[i];

        // sum of X[i] * Y[i].
        sum_XY = sum_XY + X[i] * Y[i];

        // sum of square of array elements.
        squareSum_X = squareSum_X + X[i] * X[i];
        squareSum_Y = squareSum_Y + Y[i] * Y[i];
    }

    // use formula for calculating correlation coefficient.
    float corr = (float)(n * sum_XY - sum_X * sum_Y)
                  / sqrt((n * squareSum_X - sum_X * sum_X)
                      * (n * squareSum_Y - sum_Y * sum_Y));

    return corr;
}

// Driver function
int main()
{

    int X[] = {15, 18, 21, 24, 27};
    int Y[] = {25, 25, 27, 31, 32};

    //Find the size of array.
    int n = sizeof(X)/sizeof(X[0]);

    //Function call to correlationCoefficient.
    cout<<correlationCoefficient(X, Y, n);

    return 0;
}
```

Java

```
// JAVA Program to find correlation coefficient
import java.math.*;

class GFG {

    // function that returns correlation coefficient.
    static float correlationCoefficient(int X[],
                                        int Y[], int n)
    {

        int sum_X = 0, sum_Y = 0, sum_XY = 0;
        int squareSum_X = 0, squareSum_Y = 0;

        for (int i = 0; i < n; i++)
        {
            // sum of elements of array X.
            sum_X = sum_X + X[i];

            // sum of elements of array Y.
            sum_Y = sum_Y + Y[i];

            // sum of X[i] * Y[i].
            sum_XY = sum_XY + X[i] * Y[i];

            // sum of square of array elements.
            squareSum_X = squareSum_X + X[i] * X[i];
            squareSum_Y = squareSum_Y + Y[i] * Y[i];
        }

        // use formula for calculating correlation
        // coefficient.
        float corr = (float)(n * sum_XY - sum_X * sum_Y)/
                     (float)(Math.sqrt((n * squareSum_X -
                     sum_X * sum_X) * (n * squareSum_Y -
                     sum_Y * sum_Y)));

        return corr;
    }

    // Driver function
    public static void main(String args[])
    {

        int X[] = {15, 18, 21, 24, 27};
        int Y[] = {25, 25, 27, 31, 32};

        // Find the size of array.
        int n = X.length;

        // Function call to correlationCoefficient.
        System.out.printf("%6f",
                 correlationCoefficient(X, Y, n));


    }
}

/*This code is contributed by Nikita Tiwari.*/
```

```
0.953463
```

Python

```
# Python Program to find correlation coefficient.
import math

# function that returns correlation coefficient.
def correlationCoefficient(X, Y, n) :
    sum_X = 0
    sum_Y = 0
    sum_XY = 0
    squareSum_X = 0
    squareSum_Y = 0


    i = 0
    while i < n :
        # sum of elements of array X.
        sum_X = sum_X + X[i]

        # sum of elements of array Y.
        sum_Y = sum_Y + Y[i]
        # sum of X[i] * Y[i].
        sum_XY = sum_XY + X[i] * Y[i]

        # sum of square of array elements.
        squareSum_X = squareSum_X + X[i] * X[i]
        squareSum_Y = squareSum_Y + Y[i] * Y[i]

        i = i + 1

    # use formula for calculating correlation
    # coefficient.
    corr = (float)(n * sum_XY - sum_X * sum_Y)/
           (float)(math.sqrt((n * squareSum_X -
           sum_X * sum_X)* (n * squareSum_Y -
           sum_Y * sum_Y)))
    return corr

# Driver function
X = [15, 18, 21, 24, 27]
Y = [25, 25, 27, 31, 32]

# Find the size of array.
n = len(X)

# Function call to correlationCoefficient.
print ('{0:.6f}'.format(correlationCoefficient(X, Y, n)))

# This code is contributed by Nikita Tiwari.
```

C#

```
// C# Program to find correlation coefficient
using System;

class GFG {

    // function that returns correlation coefficient.
    static float correlationCoefficient(int []X, int []Y,
                                                   int n)
    {
        int sum_X = 0, sum_Y = 0, sum_XY = 0;
        int squareSum_X = 0, squareSum_Y = 0;

        for (int i = 0; i < n; i++)
        {
            // sum of elements of array X.
            sum_X = sum_X + X[i];

            // sum of elements of array Y.
            sum_Y = sum_Y + Y[i];

            // sum of X[i] * Y[i].
            sum_XY = sum_XY + X[i] * Y[i];

            // sum of square of array elements.
            squareSum_X = squareSum_X + X[i] * X[i];
            squareSum_Y = squareSum_Y + Y[i] * Y[i];
        }

        // use formula for calculating correlation
        // coefficient.
        float corr = (float)(n * sum_XY - sum_X * sum_Y)/
                     (float)(Math.Sqrt((n * squareSum_X -
                     sum_X * sum_X) * (n * squareSum_Y -
                     sum_Y * sum_Y)));

        return corr;
    }
    // Driver function
    public static void Main()
    {

        int []X = {15, 18, 21, 24, 27};
        int []Y = {25, 25, 27, 31, 32};

        // Find the size of array.
        int n = X.Length;

        // Function call to correlationCoefficient.
        Console.Write(Math.Round(correlationCoefficient(X, Y, n) *
                                            1000000.0)/1000000.0);


    }
}

//This code is contributed by Anant Agarwal.
```

PHP

```
<?php
// PHP Program to find
// correlation coefficient

// function that returns
// correlation coefficient.
function correlationCoefficient($X, $Y, $n)
{
    $sum_X = 0;$sum_Y = 0; $sum_XY = 0;
    $squareSum_X = 0; $squareSum_Y = 0;

    for ($i = 0; $i < $n; $i++)
    {
        // sum of elements of array X.
        $sum_X = $sum_X + $X[$i];

        // sum of elements of array Y.
        $sum_Y = $sum_Y + $Y[$i];

        // sum of X[i] * Y[i].
        $sum_XY = $sum_XY + $X[$i] * $Y[$i];

        // sum of square of array elements.
        $squareSum_X = $squareSum_X +
                       $X[$i] * $X[$i];
        $squareSum_Y = $squareSum_Y +
                       $Y[$i] * $Y[$i];
    }

    // use formula for calculating
    // correlation coefficient.
    $corr = (float)($n * $sum_XY - $sum_X * $sum_Y) /
         sqrt(($n * $squareSum_X - $sum_X * $sum_X) *
              ($n * $squareSum_Y - $sum_Y * $sum_Y));

    return $corr;
}

// Driver Code
$X = array (15, 18, 21, 24, 27);
$Y = array (25, 25, 27, 31, 32);

//Find the size of array.
$n = sizeof($X);
```

```
//Function call to
// correlationCoefficient.
echo correlationCoefficient($X, $Y, $n);

// This code is contributed by aj_36
?>
```

Output :

```
0.953463
```

Reference –
Correlation coefficient – Wikipedia

This article is contributed by **Dharmendra Kumar**. If you like GeeksforGeeks and would like to contribute, you can also write an article using contribute.geeksforgeeks.org or mail your article to contribute@geeksforgeeks.org. See your article appearing on the GeeksforGeeks main page and help other Geeks.

Please write comments if you find anything incorrect, or you want to share more information about the topic discussed above.

Recommended Posts:

Program for Spearman's Rank Correlation
Program for Coefficient of variation
Binomial Coefficient | DP-9
Sum of product of r and rth Binomial Coefficient (r * nCr)
Maximum binomial coefficient term value
Clustering Coefficient in Graph Theory
Space and time efficient Binomial Coefficient
Replace the maximum element in the array by coefficient of range
Eggs dropping puzzle (Binomial Coefficient and Binary Search Solution)
Program to find sum of 1 + x/2! + x^2/3! +...+x^n/(n+1)!
Fibonomial coefficient and Fibonomial triangle
Program to find the value of sin(nΘ)
Program to Find the value of cos(nΘ)
Program to find the value of tan(nΘ)
Program to find sum of the given sequence

Improved By : jit_t

Article Tags : Mathematical statistical-algorithms

Practice Tags : Mathematical

👍

2

□ To-do □ Done

2

Based on **2** vote(s)

Feedback/ Suggest Improvement Add Notes Improve Article

Please write to us at contribute@geeksforgeeks.org to report any issue with the above content.

Writing code in comment? Please use ide.geeksforgeeks.org, generate link and share the link here.

Load Comments

A computer science portal for geeks

5th Floor, A-118,
Sector-136, Noida, Uttar Pradesh - 201305
feedback@geeksforgeeks.org

COMPANY
About Us
Careers
Privacy Policy
Contact Us

LEARN
Algorithms
Data Structures
Languages
CS Subjects
Video Tutorials

PRACTICE
Courses
Company-wise
Topic-wise
How to begin?

CONTRIBUTE
Write an Article
Write Interview Experience
Internships
Videos

@geeksforgeeks, Some rights reserved