IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 3, March 2022

DOI: 10.17148/IJARCCE.2022.11357

STATISTICS PROBLEMS USING MAXIMA SOFTWARE

Varalaxmi T. Shedole[1] and Indrani Y.R.L[2]

[1]Assistant Professor, Department of Statistics, Nrupathunga University, Bengaluru , Karnataka.

[2]Assistant Professor, Department of Mathematics, Nrupathunga University, Bengaluru, Karnataka.

Abstract: Maxima is a computer algebra system software which is Free Open Source Software (FOSS). In this paper, we solved statistics problems on correlation coefficient and Spearmn's rank correlation using maxima where speed matters for numerical computation. With easy and simple commands it reduces the time taken for tedious calculations of lengthy problems and helps in obtaining quick solutions for similar statistical problems. Overall the maxima software is a friendly software which is easily accessible. In future, it will be definitely a stepping stone for students, teachers and for researchers in the field of statistics. We solved several examples to exploit the capability of maxima by its many commands for computations in statistics.

Key words: Correlation coefficient, Spearman's rank correlation coefficient, Maxima, FOSS.

1. INTRODUCTION

Maxima provides an environment for statistical analysis in which symbolic computation can be easily used together with all the facilities [1]. The aim of this paper is to examine the utility of Maxima a FOSS software applications in the office, class room and for research purpose. Maxima is applicable in getting numerical results using exact fractions, integers and variable precision [2]. Currently we use R programming, which is a robust statistical package. Maxima software provides an integrated environment for data manipulation, statistical analysis, graphical display and more. A computer algebra system (CAS) is a software program that can perform symbolic computations such as ratios, numbers with arbitrary numerical precision, equations, integrals, and well-known mathematical constants [4].

Here we applied these software commands to solve statistics problems on correlation coefficient and rank correlation coefficient.

Let x and y be the two random variables, Prof. Karl Pearson discovered the linear relationship between these two random variables [5], which is known as product moment correlation and is denoted by $r(x,y)$. The correlation between these two random variables (S.C. Gupta and V.K. Kapoor). is defined as $r = r(x,y) = \dfrac{\mathrm{cov}(x,y)}{\sigma_x \sigma_y}$,

where $r(x,y) = \dfrac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\bar{y}$, $\sigma_x = \sqrt{\dfrac{1}{n}\sum_{i=1}^{n} x_i^{\,2} - \left(\bar{x}\right)^2}$, $\sigma_y = \sqrt{\dfrac{1}{n}\sum_{i=1}^{n} y_i^{\,2} - \left(\bar{y}\right)^2}$,

i=1,2,3,.....n.

Let $x_1, x_2, \ldots\ldots x_n$ and $y_1, y_2, \ldots\ldots y_n$ be 'n' individuals, then Spearman's rank correlation coefficient is given as , $\rho = 1 - \dfrac{6\sum_{i=1}^{n} di^{\,2}}{n(n^2-1)}$, where $d_i = x_i - y_i \quad i = 1,2,\ldots\ldots n$.

2. PRELIMINARIES

Using the simple commands of Maxima software, the statistics problems are solved on correlation coefficient and Spearman's rank correlation coefficient.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 3, March 2022

DOI: 10.17148/IJARCCE.2022.11357

Example 2.1: The following table gives number of hours prepared for an examination and the marks of 10 students. Find the correlation between them.

Student number	1	2	3	4	5	6	7	8	9	10
Number of hours prepared	4	9	10	14	4	7	12	22	1	17
Marks obtained	31	58	65	73	37	44	60	91	21	84

Solution:

Let x be number of hours prepared for examination
Let y be number of hours prepared for examination

x_i	y_i	x_i^2	y_i^2	$x_i y_i$
4	31	16	961	124
9	58	81	3364	522
10	65	100	4225	650
14	73	196	5329	1022
4	47	16	1369	148
7	44	49	1936	308
12	60	144	3600	720
22	91	484	8281	2002
1	21	1	441	21
17	84	289	7056	1428
$\sum_{i=1}^{n} x_i = 100$	$\sum_{i=1}^{n} y_i = 564$	$\sum_{i=1}^{n} x_i^2 = 1376$	$\sum_{i=1}^{n} y_i^2 = 36562$	$\sum_{i=1}^{n} x_i y_i = 6945$

n=10

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{100}{10} = 10 \qquad \bar{y} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{564}{10} = 56.4$$

$$\sigma_x = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i - (\bar{x})^2} = \sqrt{\frac{1376}{10} - (10)^2} = 6.1319$$

$$\sigma_y = \sqrt{\frac{1}{n}\sum_{i=1}^{n} y_i - (\bar{y})^2} = \sqrt{\frac{36562}{10} - (56.4)^2} = 21.8$$

$$\mathrm{cov}(x, y) = \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}.\bar{y} = \frac{6945}{10} - 10 \times 6.4 = 130.5$$

$$r_{xy} = \frac{\mathrm{cov}(x, y)}{\sigma_x \sigma_y} = \frac{130.5}{(6.1319)(21.8)} = 0.9762$$

Therefore there is a highly positive correlation between x and y.
Hence if a student works more, he will get more marks

Maxima Program

```
kill(all)$
d1:[1,2,3,4,5,6,7,8,9,10];
d2:[4,9,10,14,4,7,12,22,1,17]$
d3:[31,58,65,73,37,44,60,91,21,84]$
d4:d2^2$
d5:d3^2$
```

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 3, March 2022

DOI: 10.17148/IJARCCE.2022.11357

```
d6:d2*d3$
d:[d2,d3,d4,d5,d6];
n:length(d2)$
s:[0,0,0,0,0]$
m:length(s)$
for j:1 thru m do (
   for i:1 thru n do (
   s[j]:s[j]+d[j][i]),
disp(s[j]))$
x1:s[1]/n;
y1:float(s[2]/n);
σx:float(sqrt((1/n)*s[3]-x1^2));
σy:float(sqrt((1/n)*s[4]-y1^2));
cxy:(1/n)*s[5]-x1*y1;
rxy:cxy/(σx*σy);
fpprintprec:6$
```

Output

```
(%o1)  [1,2,3,4,5,6,7,8,9,10]
(%o7)  [[4,9,10,14,4,7,12,22,1,17],[31,58,65,73,37,44,60,91,21,84],[16,
81,100,196,16,49,144,484,1,289],[961,3364,4225,5329,1369,1936,3600
,8281,441,7056],[124,522,650,1022,
148,308,720,2002,21,1428]]
100
564
1376
36562
6945
(%o12)  10
(%o13)  56.4
(%o14)  6.13188
(%o15)  21.8
(%o16)  130.5
(%o17)  0.976248
```

Example 2.2 : Students were participated in a music competition and three judges have given the ranks. Obtain Spearman's rank correlation coefficient and compare them.

Ranks of Judge 1	1	5	6	10	3	2	9	4	11	7	12	8
Ranks of Judge 2	3	8	5	7	4	6	11	1	9	2	10	12
Ranks of Judge 3	4	6	9	8	11	3	12	2	10	5	1	7

Solution: Let us denote x, y and z are ranks given by three judges in music competition.
Here n = 12

x	y	z	d_{xy}	d_{xz}	d_{yz}	d_{xy}^2	d_{xz}^2	d_{yz}^2
1	3	4	-2	-3	-1	4	9	1
5	8	6	-3	-1	2	9	1	4
6	5	9	1	-3	-4	1	9	16
10	7	8	3	2	-1	9	4	1
3	4	11	-1	-8	-7	1	64	49
2	6	3	-4	-1	3	16	1	9
9	11	12	-2	-3	-1	4	9	1
4	1	2	3	2	-1	9	4	1
11	9	10	2	1	-1	4	1	1
7	2	5	5	2	-3	25	4	9
12	10	1	2	11	9	4	121	81
8	12	7	-4	1	5	16	1	25
						$\sum d_{xy}^2$ =102	$\sum d_{xy}^2$ =228	$\sum d_{xy}^2$ =198

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 3, March 2022

DOI: 10.17148/IJARCCE.2022.11357

Rank correlation coefficient of ranks given by judges 1 and 2

$\rho_{12} = 1 - \frac{6\sum d_{xy}^2}{n(n^2-1)} = 1 - \frac{6*102}{12(12^2-1)} = 0.6434$

Rank correlation coefficient of ranks given by judges 1 and 3

$\rho_{13} = 1 - \frac{6\sum d_{xy}^2}{n(n^2-1)} = 1 - \frac{6*228}{12(12^2-1)} = 0.2028$

Rank correlation coefficient of ranks given by judges 2 and 3

$\rho_{23} = 1 - \frac{6\sum d_{xy}^2}{n(n^2-1)} = 1 - \frac{6*198}{12(12^2-1)} = 0.3077$

The rank correlation coefficient is high for the ranks of 1 and 2 which is 0.6434. i.e., their judgements are near.

Maxima Program:

```
(%i16)  kill(all)$
        x:[1,5,6,10,3,2,9,4,11,7,12,8]$
        y:[3,8,5,7,4,6,11,1,9,2,10,12]$
        z:[4,6,9,8,11,3,12,2,10,5,1,7]$
        dxy1:x-y;
        dxz1:x-z;
        dyz1:y-z;
        d1:(x-y)^2$
        d2:(x-z)^2$
        d3:(y-z)^2$
        d:[d1,d2,d3];
        n:length(x);
        s:[0,0,0]$
        m:length(s)$
        for j:1 thru m do(
           for i:1 thru n do (
           s[j]:s[j]+d[j][i]),
        disp(s[j]))$
        for i:1 thru m do (
           R[i]:float(1-((6*s[i])/(n*(n^2-1)))),
        disp(R[i]))$
        fpprintprec:6$
```

Output

```
(%o4)  [-2,-3,1,3,-1,-4,-2,3,2,5,2,-4]
(%o5)  [-3,-1,-3,2,-8,-1,-3,2,1,2,11,1]
(%o6)  [-1,2,-4,-1,-7,3,-1,-1,-1,-3,9,5]
(%o10) [[4,9,1,9,1,16,4,9,4,25,4,16],[9,1,9,4,64,1,9,4,1,4,121,1],[1,4,16,1,49,9,1,1,1,9,81,25]]
(%o11) 12
102
228
198
0.643357
0.202797
0.307692
```

3. THE DIFFRERENT PROBLEMS ARE SOLVED WITH OUPUT RESULTS

Different types of problems are solved using Maxima FOSS and resluting otputs are obtained very easily just by using its simple and easy commands.

Example 3.1: The marks obtained by 10 students in Mathematics and Statistics are given below. Find the correlation coefficient between the two subjects.

Marks in Mathematics	75	30	60	80	53	35	15	40	38	48
Marks in Statistics	85	45	54	91	58	63	35	43	45	44

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 3, March 2022

DOI: 10.17148/IJARCCE.2022.11357

Solution :

Ranks of Mathematics	1	2	6	9	11	15	10	8	4	7	5	14	13	12	3
Ranks of Statistics	10	7	8	11	9	13	15	1	6	3	4	12	14	5	2

Output

(%o1) [75,30,60,80,53,35,15,40,38,
48]
(%o2) [85,45,54,91,58,63,35,43,45,
44]
(%o8) [[22,-23,7,27,0,-18,-38,-13,-

15,-5],[22,-18,-9,28,-5,0,-28,-20,
-18,-19],[484,529,49,729,0,324,
1444,169,225,25],[484,324,81,784,
25,0,784,400,324,361],[484,414,-63
,756,0,0,1064,260,270,95]]
-56
-67
3978
3567
3280
(%o13) -5.6
(%o14) -6.7
(%o15) 19.1426
(%o16) 17.6581
(%o17) 290.48
(%o18) 0.859349

Example 3.2: The ranks of 15 students in Mathematics and Statistics are given below. Obtain rank correlation coefficient between them.

Temperature	57	42	40	38	42	45	42	44	40	46	44	43
Germination time	10	26	30	41	29	27	27	19	18	19	31	29

Solution :

Output

(%o3) [-9,-5,-2,-2,2,2,-5,7,-2,4,1,2,-1,7,1]
(%o5) [[81,25,4,4,4,4,25,49,4,16,1,4,1,49,1]]
272
0.514286

Example 3.3: The following table gives the soil temperature and the germination time at various places. Calculate the co-efficient of correlation and interpret the value.

Solution :

(%o1) [57,42,40,38,42,45,42,44,40,46,44,43]
(%o2) [10,26,30,41,29,27,27,19,18,19,31,29]
(%o3) 44
(%o4) 26
(%o10) [[13,-2,-4,-6,-2,1,-2,0,-4,2,0,-1],[-16,0,4,15,3,1,1,-7,-8,-7,5,3],[169,4,16,36,4,1,4,0,16,4,0,1],
[256,0,16,225,9,1,1,49,64,49,25,9],[-208,0,-16,-90,-6,1,-2,0,32,-14,0,-3]]
-5
-6
255

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 3, March 2022

DOI: 10.17148/IJARCCE.2022.11357

704
-306
(%o15) -0.416667
(%o16) -0.5
(%o17) 4.5909
(%o18) 7.64308
(%o19) -25.7083
(%o20) -0.732668

Example 3.4: With the following data in 6 cities, calculate the coefficient of correlation by Pearson's method between the density of population and death rate.

Cities	A	B	C	D	E	F
Area in Sq.miles	150	180	100	60	120	80
Population in '000	30	90	40	42	72	24
Number of deaths	300	1440	560	840	1224	312

Solution :

The output using maxima FOSS

[150,180,100,60,120,80]
(%o2) [30000,90000,40000,42000,72000,24000]
(%o3) [300,1440,560,840,1224,312]
(%o4) [200,500,400,700,600,300]
(%o5) [10,16,14,20,17,13]
(%o6) 450.0
(%o7) 15.0
(%o8) [-250.0,50.0,-50.0,250.0,150.0,-150.0]
(%o9) [-5.0,1.0,-1.0,5.0,2.0,-2.0]
(%o10) [62500.0,2500.0,2500.0,62500.0,22500.0,22500.0]
(%o11) [25.0,1.0,1.0,25.0,4.0,4.0]
(%o12) [1250.0,50.0,50.0,1250.0,300.0,300.0]
(%o13) [[-250.0,50.0,-50.0,250.0,150.0,-150.0],[-5.0,1.0,-1.0,5.0,2.0,-2.0],
[62500.0,2500.0,2500.0,62500.0,22500.0,22500.0],[25.0,1.0,1.0,25.0,4.0,4.0],[1250.0,50.0,50.0,1250.0,300.0,300.0]]
0.0
0.0
1.75·10^5
60.0
3200.0
(%o18) 0.987541

4. RESULTS AND CONCLUSION:

- These software packages relieve the risk of obtaining incorrect results from algebraic and numerical manipulations.
- Here this software is used for solving problems on numerical calculations of statistics problems, which can also be used for symbolic calculations and graphical solutions etc.
- When speed matters, this software gives the quick results in case of extreme values used mainly for the calculations of statistical problems which is helpful in many sectors.

REFERENCES

[1] Giovanni L. Lo Magno (2015). More power through symbolic computation: Extending Stata by using the Maxima computer algebra system The Stata Journal 15, Number 1, pp. 45–76.
[2] Darren D. Wick (2009). Free and open-source software for mathematics and education, proceedings pdf pp. 1-5.

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 3, March 2022

DOI: 10.17148/IJARCCE.2022.11357

[3] D Afriza[1], C M Zbainur[1] and A Veloo[2](2018). The use of the wxMaxima linear algebra module on Gauss elimination lession for mathematics education students. IOP publishing, Journal of Physics: Conf. Series 1088/1742-6596.

[4] Grabmeier, J., E. Kaltofen, and V. Weispfenning, eds.(2003). Computer Algebra Handbook: Foundations, Applications, Systems.

[5] S.C. Gupta and V.K. Kapoor (Reprint 2016) A Text Book on Fundamentals of Mathematical Statistics. Published by SULTAN CHAND & SONS, ISBN: 978-81-8054-528-3.

[6] H.V.Geetha[1],T.G.Sudha[2] and Harshini Srinivas[3] (2017) Solution of Wave Equation by the Method of Separation of Variables Using Foss Tools Maxima, International Journal of Pure and Applied Mathematics, Volume 117 No. 14,167-174.

[7] D.V.L.N Jogiraju, C. Srikala and Dr. L.P.Rajkumar (2018) A text book on Statistical Methods, by Kalyani Publishers, ISBN :978-93-272-8189-7.