Submitted by-Aditya Gautam

Roll No-12

M.sc. 3rd semester

Date of Assignment-03/12/2020

Date of Submission-11/12/2020

**Experiment No-08**

**Topic**-DISCRIMINANT ANALYSIS AND MAHALANOBIS D2 STATISTIC.

**Problem-** The following table shows the marks obtained in Mathematics(x) and marks obtained in statistics (y) by two batches of students.

|  |  |
| --- | --- |
| **BATCH-1** | |
| x | y |
| 12 | 34 |
| 45 | 56 |
| 44 | 35 |
| 52 | 63 |
| 8 | 32 |
| 39 | 48 |
| 71 | 84 |
| 38 | 57 |
| 38 | 51 |
| 47 | 62 |

|  |  |
| --- | --- |
| **BATCH-2** | |
| x | Y |
| 56 | 54 |
| 67 | 66 |
| 49 | 72 |
| 89 | 97 |
| 58 | 76 |
| 53 | 32 |
| 56 | 81 |
| 78 | 98 |
| 64 | 78 |
| 58 | 40 |

Compute the MAHALANOBIS- D2  statistic and hence performD2 test for testing the equality of marks obtained by the two groups. Also, classify to which batch a student scoring 52 in mathematics and 70 in statistic will belong.

**Theory-**

The Fisher’s discriminant function is given by

=(/ S-1pooledX̰ X̰=

Where,  =((1,1,...........(1) **;** (1=1 i=1,2,......p

=((2,2,...........(2) **;** (2=2 i=1,2,......p

Spooled=

Where, S1=((Sij)1),(Sij)1)=Xik)1-(i)1}{(Xjk)1 - (j)1}

S2=((Sij)2),(Sij)2)=Xik)2-(i)2}{(Xjk)2 - (j)2}

The function is a function which maximally separates the two populations and the maximum separation in the two sample from the population is D2=(/ S-1pooled

Which is the MAHALANOBIS D2 STATISTIC .

and

Here, are two samples from the multivariate normal populations Np(µ̰1,, Np(µ̰2,respectively. Further, it is assumed that two population have the same variance covariance matrix.

Here, we are to test the hypothesis H0:(µ̰1 - µ̰2)2=0̰ against H1:(µ̰1 - µ̰2)2≠0̰

i.e. the hypothesis of equality of the two population means . It can be accomplished from the basis of the D2 statistic for the D2 test is D2p,n1+n2-p-1

Conclusions are drawn accordingly

For allocating the observation X̰= into one of the two groups , the allocation rule based on is –

1. Allocate X̰0 to Np(µ̰1, (group 1) if 0 = (/ S-1pooledX̰0 ≥ =
2. Allocate X̰0 to Np(µ̰2, (group 2) if 0 **=** (/ S-1pooledX̰0 <

**Calculation-**

The R-programming to obtain the solution for the given problem-

x1=c(12,45,44,52,8,39,71,38,38,47,34,56,35,63,32,48,84,57,51,62)

dim(x1)=c(10,2)

x2=c(56,67,49,89,58,53,56,78,64,58,54,66,72,97,76,32,81,98,78,40)

dim(x2)=c(10,2)

dim(x2)

mean1=mat.or.vec(2,1)

mean2=mat.or.vec(2,1)

for(i in 1:2){

mean1[i]=mean(x1[,i])

mean2[i]=mean(x2[,i])}

mean=array(c(mean1-mean2),dim=c(2,1))

mean

n1=10

n2=10

var11=mat.or.vec(2,1)

var12=mat.or.vec(2,1)

var21=mat.or.vec(2,1)

var22=mat.or.vec(2,1)

for(i in 1:2){

var11[i]=cov(x1[,1],x1[,i])\*((n1-1)/(n1+n2-2))

var12[i]=cov(x2[,1],x2[,i])\*((n2-1)/(n1+n2-2))}

for(i in 1:2){

var21[i]=cov(x1[,2],x1[,i])\*((n1-1)/(n1+n2-2))

var22[i]=cov(x2[,2],x2[,i])\*((n2-1)/(n1+n2-2))}

s1=c(var11,var21)

s1

dim(s1)=c(2,2)

dim(s1)

s2=c(var12,var22)

s2

dim(s2)=c(2,2)

dim(s2)

s\_p=s1+s2

s\_p

D2=t(mean)%\*%solve(s\_p)%\*%mean

D2

p=2

cal\_value=((n1+n2-p-1)/(p\*(n1+n2-2)))\*((n1\*n2)/(n1+n2))\*D2

cal\_value

tab\_value=qf(0.95,2,17,0)

tab\_value

x0=array(c(52,70),dim=c(2,1))

x0

y0=t(mean)%\*%solve(s\_p)%\*%x0

y0

m=D2/2

m

**Conclusion-**

The MAHALANOBIS D2- STATISTIC is 2.333933 . Since the calculated value (i.e. 5.510676) is more than the tabulated value (i.e. 3.591531) of F we reject our null hypothesis at 5% level of significance and conclude that the equality of marks obtained by the two groups are significantly different.

And a student scoring 52 in Mathematics and 70 in statistics will belong to batch 2 since

0(=- 4.515759) <(=1.166967)