TP report: "FACTORIAL ANALYSIS OF CORRESPONDENCE"

BRAHMIA ABDELBACET

April 11, 2021

1 introduction

1.1 Definition of Factorial Analysis Of Correspondence :

is a multivariate technique that may be applied to any type of data . It detects associations and oppositions existing between subjects and objects.

1.2 The objective of Principal Component Analysis :

- The aim of correspondence analysis is to illustrate the relationships between two categorial variables.
- This method attempts to retain as far as possible the distances that exist between the categories of a variable with respect to its description by the respective other variable when generating a reduced solution.

1.3 presentation of our inputs

1.3.1 Presentation of Khi-2 table

Generally, the khi-2 table give precisely the probabilities P (X¿ t) for different values of t and different degrees of freedom.

1.3.2 Presentation of observed data table

A raw data table that we seek to prove the similarity or difference between the variables it contains.

Code Matlab :

```
donnees_observees=[36 34 14 16]
```

1.4 The steps followed

1.4.1 calculate the table of theoretical values

in our case we only have one line , so we calculate it with the following steps

Code Matlab :

```
n=sum(donnees_observees)
val=n/w;
donnees_theoriques(1:4)=val
```

in the case that we have multi line we calculate the table with the following steps

Code Matlab :

```
somme1= sum(donnees_observees_2,1)
somme2= sum(donnees_observees_2,2)
somme3=sum(somme2,1)
for i=1:w
    for j=1:h
        donnees_theoriques(i,j)= (somme1(j)*somme2(i))/somme3;
    end
end
```

1.4.2 X-calculated

we calculate the value of X-calculated with the formula $\sqrt{\frac{(\Delta)^2}{Te}}$
suchas $\Delta = To - Te$

Code Matlab :

```
x_calculer=0;
for i=1:h
    for j=1:w
        x_calculer=x_calculer+( (donnees_observees(i,j) -
        donnees_theoriques(i,j) ).^2 /donnees_theoriques(i,j));
    end
end
```

1.4.3 degree of liberty

Code Matlab :

```
ddl=(w-1)*(h-1)
```

1.4.4 X-theoretical

In this example we have used significance level = 0.05
then we look for the value in the intersection of our significance level value with the value closest to our dll value

Code Matlab :

```
for i=1:size(p,2)
    if p(i)==niveau_de_signification
        val2=i+1;
        break
    end
end

KHI_2(:,1)
for i=1:size(KHI_2(:,1),2)
    if KHI_2(i,1)>=ddl
        val3=i
        break
    end
end

x_theorique=KHI_2(val3,val2)
```

1.4.5 null hypothesis

deduce the rejection / non-rejection of the null hypothesis and deduce that the rows and columns are independent / dependent .

the contingency coefficient If the link exists, then we calculate the contingency coefficient and we deduce whether the link is strong or weak .
contingency coefficient=$\sqrt{\frac{Xc^2}{n+Xc^2}}$

- Si La valeur de ce coefficient est proche de 1, alors la liaison est forte
- S'il est proche de 0, la liaison est faible.

```
if (x_calculer > x_theorique )

    disp('null hypothesis rejected, So there is a link between
        the factors and the choice')
    le_coefficient_de_contingence=  ( (x_calculer).^2/(n+
                                (x_calculer).^2 )  ).^(1/2)

        if le_coefficient_de_contingence<0.5
            disp("The value is close to 0 , the bond is weak.")
        else
            disp("The value is close to 1, the bond is strong.")
        end

else
    disp('null hypothesis accepted, So there is no link between
        the factors and the choice')
end
```

1.4.6 Diagram the response

Code Matlab :

```
x = 0:0.1:16;
y = chi2pdf(x,3);
x2= x_theorique


figure;
plot(x2,y,'g*')
hold on
plot(x,y,'b-o')
xlabel('Observation', 'Color', 'b')
ylabel('Probability Density', 'Color', 'b')
hold on
```

In our example, we have the following diagram as result:

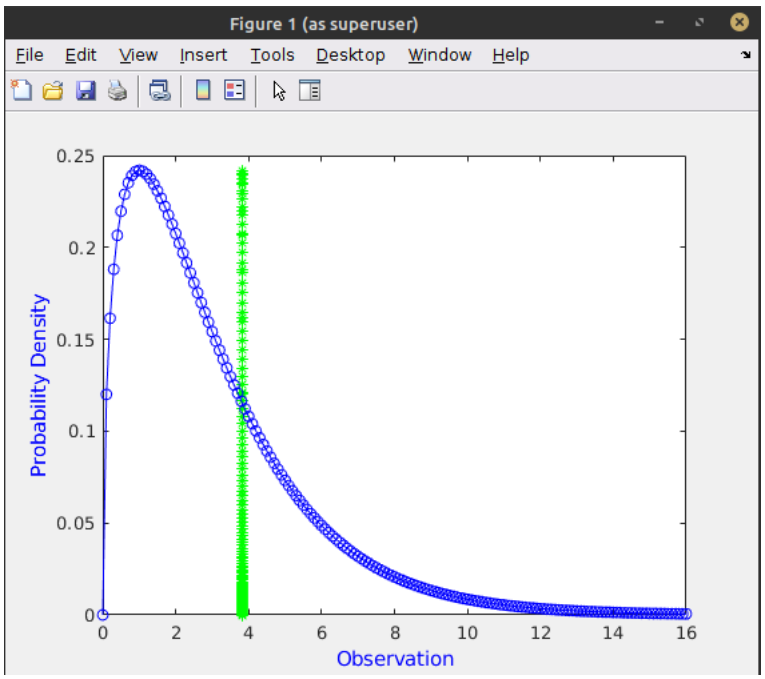

Figure 1: diagram

1.5 conclusion

Our result:

- ddl = 0
- donnees theoriques = 100 100 100 100
- x calculer = 40.9600
- x theorique = 3.8415
- null hypothesis rejected, So there is a link between the factors and the choice
- le coefficient de contingence = 0.9715
- The value is close to 1, the bond is strong.