1. Introduction

Welcome to the series of E-learning modules on Practical Problems on Association of Attributes. Here we have worked out the problems on finding the missing frequencies in two and three attributes, consistency of the given data, frequencies of positive and negative classes, testing for the independence of events and obtaining the Yule's coefficient of association and drawing inference based on this.

By the end of this session, you will be able to

- Find the missing frequencies
- Verify the consistency of then given data
- Verify the association between the attributes to know if they are dependent or independent
- Find Yule's coefficient of association

Before we start practical exercises, let us have a glance over the different terms and measures used in this practical.
The number of units in different classes is called "class frequencies". Thus if number of blind and deaf people is twenty, the frequency of class AB is twenty. Class frequencies are denoted by enclosing the class symbols brackets. Thus (AB) would represent the frequency of the class AB.

2. Consistency of Data, Types of Classes and Independent Attributes

Now let us discuss about the consistency of the data
It is obvious that in statistics of attributes when frequencies of various classes are counted, no class frequency can be negative. If any class frequency is negative the data are said to be inconsistent.
In order to test whether a set of figures is consistent, various class frequencies should be found and if no class frequency is negative apparently the data are consistent.

Now let us discuss the different types of classes.
The classes which represent the presence of an attribute or attributes are called positive classes. The classes which represent the absence of an attribute or attributes are called negative classes. The classes in which one attribute is present and the other is absent are called pairs of contraries. Thus
N, A, B and AB are positive classes
Alpha and beta are negative classes
A beta and alpha beta are pairs of contrary.

When can we say that the two attributes are independent?
Two attributes A and B are said to be independent if there exists no relationship of any kind between them. If A and B are independent we would expect

i. The same proportion of A's amongst B's as amongst beta's and
ii. The proportion of B's amongst A's is same as that amongst the alphas.

One of the ways of calculating it is that if expected frequency and given frequencies are same, we can say the attributes are independent.

Yule's coefficient of association is given by,
Q is equal to (AB) into (alpha beta) minus (A beta) into (alpha B) divided by (AB) into (alpha beta) plus (A beta) into (alpha B)

3. Exercise(Part 1)

Exercise one
Given the following ultimate class frequencies, find the frequencies of the positive and negative classes and total number of observations
(AB) is equal to two hundred and fifty, (A beta) is equal to one hundred and twenty, (alpha B) is equal to two hundred, and (alpha beta) is equal to seventy.

Solution:
We know that,
N is equal to (A) plus (α) is equal to (AB) plus (A beta) plus (alpha B) plus (alpha beta) is equal to two hundred and fifty plus one hundred and twenty plus two hundred plus seventy is equal to six hundred and forty.
Now let us find the frequencies of positive classes.
(A) is equal to (AB) plus (A beta) is equal to two hundred and fifty plus one hundred and twenty is equal to three hundred and seventy
(B) is equal to (AB) plus (alpha B) is equal to two hundred and fifty plus two hundred is equal to four hundred and fifty
Frequencies of negative classes are given by,
(alpha) is equal to(alpha B) plus (alpha beta) is equal to two hundred plus seventy is equal to two hundred and seventy
(beta) is equal to (A beta) plus (alpha beta) is equal to one hundred and twenty plus seventy is equal to one hundred and ninety

Exercise 2
An investigation of twenty three thousand seven hundred and thirteen households was made in an urban and rural mixed locality. Of these one thousand six hundred and eighteen were farmers, two thousand fifteen were well-to-do and seven hundred and seventy families were having at least one graduate. Of these graduate families, three hundred and thirty five were those of farmers and four hundred and twenty eight were well-to-do, also five hundred and eighty seven well-to-do families were those of farmers and out of them one hundred and fifty six were having at least one of their family members as graduate. Obtain all ultimate class frequencies.

Solution
Let the attribute ‘farming’ be denoted by A, the attribute ‘well-to-do’ by B and ‘having at least one graduate’ by C.
Given N is equal to twenty three thousand seven hundred and thirteen, (A) is equal to one thousand six hundred and eighteen, (B) is equal to two thousand fifteen, (C) is equal to seven hundred and seventy
(AB) is equal to five hundred and eighty seven, (BC) is equal to four hundred and twenty eight, (AC) is equal to three hundred and thirty five, (ABC) is equal to one hundred and fifty six
For three attributes A, B, C the number of ultimate class frequencies is two cube is equal to eight, one of them being (ABC) is equal to one hundred and fifty six.

(AB gamma) is equal to (AB) minus (ABC) is equal to five hundred and eighty seven minus

one hundred and fifty six is equal to four hundred and thirty one

(A beta C) is equal to (AC) minus (ABC) is equal to three hundred and thirty five minus one hundred and fifty six is equal to one hundred and seventy nine.

(A beta gamma) is equal to (A) minus (AB) minus (AC) plus (ABC)
Is equal to one thousand six hundred and eighteen minus five hundred and eighty seven minus three hundred and thirty five plus one hundred and fifty six is equal to eight hundred and fifty two.

(alpha BC) is equal to (BC) minus (ABC) is equal to four hundred and twenty eight minus one hundred and fifty six is equal to two hundred and seventy two.

(alpha B gamma) is equal to (B) minus (AB) minus (BC) plus (ABC)
Is equal to two thousand fifteen minus five hundred and eighty seven minus four hundred and twenty eight plus one hundred and fifty six is equal to one thousand one hundred and fifty six.

(alpha beta C) is equal to (C) minus (AC) minus (BC) plus (ABC)
Is equal to seven hundred and seventy minus three hundred and thirty five minus four hundred and twenty eight plus four hundred and fifty six is equal to one hundred and sixty three

(alpha beta gamma) is equal to N minus (A) minus (B) minus (C) plus (AB) plus (AC) plus (BC) minus (ABC)
Is equal to twenty three thousand seven hundred and thirteen minus one thousand six hundred and eighteen minus two thousand fifteen minus seven hundred and seventy plus five hundred and eighty seven plus three hundred and thirty five plus four hundred and twenty eight minus one hundred and fifty six is equal to twenty thousand five hundred and four.

Exercise 3
Given the following, calculate the frequencies of the remaining classes.
(A) Is equal to fifty, (AB) is equal to thirty, (B) is equal to forty, (N) is equal to hundred
Solution
Here we need find the frequencies of the classes, (alpha), (beta), (A beta), (alpha B) and (alpha beta)
Instead of using the formula and finding as we have given only two attributes, we can write the given values in the table and find the missing values.

The numbers written in yellow are the given values and the numbers written in white are calculated using the given values by filling the table.

Figure 1

	A	**α**	**Total**
B	(AB) (30)	(αB) (10)	(B) (40)
β	(Aβ) (20)	(αβ) (40)	(β) (60)
Total	(A) (50)	(α) (50)	N (100)

The last row and last columns give the totals of the corresponding rows and columns. We are also given the grand total N in the table. Hence the other values are found by subtraction as follows.
Beta is equal to hundred minus forty which is equal to sixty
Alpha beta is equal to forty minus thirty is equal to ten
A beta is equal to fifty minus thirty is equal to twenty
Alpha beta is equal to sixty minus twenty is equal to forty and finally
Alpha is equal to ten plus forty is equal to fifty.

4. Exercise(Part 2)

Exercise 4
To investigate the association between eye-colour of husband and eye colour of wife the following data are available
Husband with light eyes and wives with not-light eyes is four hundred and fourteen
Husbands with not light eyes and wives with light eyes is two hundred and sixty
Husbands with not light eyes and wives with not- light eyes is two hundred and thirty eight
Husband with light eyes is four hundred
Do you find any inconsistency in the above data?

Solution
Let us denote -Husband with light eyes by A and Husband with not-light eyes by alpha
Wives with light eyes by B and wives with not-light eyes by beta,

The given data are,
(A beta) is equal to four hundred and fourteen, (alpha B) is equal to two hundred and sixty, (alpha beta) is equal to two hundred and thirty eight, and (A) is equal to four hundred
From above, (AB) is equal to (A) minus (A beta) is equal to four hundred minus four hundred and fourteen is equal to minus fourteen.
Thus AB has got a negative value. Hence the given data are inconsistent as it is obvious that no class-frequency occurring by counting real attribute can be negative.

Exercise 5
If in a village actually involved by anthrax, seventy percent of the goats are attacked and eighty five per cent have been inoculated with vaccine; what is the lowest percentage of the inoculated that must have been attacked.

Solution
Denoting the attribute of attack by A and the attribute of inoculation by B,
The given data are (A) is equal to seventy, (B) is equal to eighty five and N is equal to hundred .We have to find the lowest percentage of (AB)

Now according to conditions of consistence, (AB) is greater than or equal to zero and (AB) is greater than or equal to (A) plus (B) minus N is equal to seventy plus eighty five minus hundred is equal to fifty five
Hence the lowest percentage of the inoculated that must have been attacked is,
(AB) into hundred divided by B is equal to fifty five into hundred divided by eighty five is equal to sixty five per cent.

Exercise 6
In a population, of two hundred students, the number of married is eighty. Out of sixty students who failed, twenty four belonged to the married group. It is required to find out whether the attributes of marriage and failure are independent.

Solution :
Let the attribute of marriage be represented by A and failure by B, the actual value of (AB) is

equal to twenty four.
Expected value of (AB) is equal to (A) into (B) divided by (AB) is equal to (eighty) into (sixty) divided by two hundred is equal to twenty four
The actual value of (AB) and its expected values are equal. It means that attributes A and B are independent.

Exercise 7
Given (A) is equal to forty, (B) is equal to thirty, (AB) is equal to twenty, and N is equal to hundred. Study the association between A and B; alpha and beta; A and beta; alpha and B.
Solution :
We can represent the given data in the shape of a table and obtain the frequencies of the missing classes.

The last row and last columns give the totals of the corresponding rows and columns. The numbers written in white and given values and the numbers written in yellow are calculated using the given values by filling the table.

Figure 2

	A	α	Total
B	(AB) (20)	(αB) (10)	(B) (30)
B	(Aβ) (20)	(αβ) (50)	(β) (70)
Total	(A) (40)	(α) (60)	N (100)

Now we write the table to get the values of different combinations namely,
Beta is equal to hundred minus thirty, is equal to seventy
Alpha B is equal to thirty minus twenty, which is equal to ten,
A beta is equal to forty minus twenty which is equal to twenty
Alpha is equal to hundred minus forty which is equal to sixty
And finally, alpha beta is equal to sixty minus ten is equal to fifty
The above values are those which have been observed. We can now calculate expected frequencies. The expected frequency of
AB is equal to A into B divided by N
Is equal to forty into thirty divided by hundred is equal to twelve.

Alpha beta is equal to alpha into beta divided by N is equal to sixty into seventy divided by hundred is equal to forty two. A beta is equal to A in to beta divided by N. which is equal to forty into seventy divided by hundred is equal to twenty eight.
Alpha B is equal to alpha into B divided by N
Is equal to sixty into thirty divided by hundred is equal to eighteen.
Now we can easily study the association between various attributes if we lay down the actual and expected values together.

The table gives the different classes, with actual values, which are from the table and expected values which are calculated as above.

Figure 3

Class	Actual Values	Expected values
(AB)	20	12
(αB)	50	42
(Aβ)	20	28
(αβ)	10	18

From the above table it is clear that the actual values of (AB) and (alpha beta) are more than the expected values and as such A and B are positively associated. Similarly, alpha and beta are also positively associated. In case of (A beta) and (alpha B), the actual values are less than the expected values. As such A and β are negatively associated and similarly there is negative association between α and B.

5. Exercise(Part 3)

Exercise 8
In a sample of five hundred children, two hundred came from higher income group, and the rest from lower income group. The number of delinquent children in these groups respectively was twenty five and hundred. Calculate the coefficient of association between delinquency and income groups.

Solution:
Let A denote the higher income group then alpha would denote the lower income group.
Let B denote delinquent children then beta would denote non-delinquent children.
To get the frequencies of the second order, we form the following table

The last row and last columns give the totals of the corresponding rows and columns. The numbers written in white and given values and the numbers written in black are calculated using the given values by filling the table.

Figure 4

	A	α	Total
B	(AB) (25)	(αB) (100)	(B) (125)
β	(Aβ) (175)	(αβ) (200)	(β) (375)
Total	(A) (200)	(α) (300)	N (500)

B is equal to twenty five plus hundred which is equal to one hundred and twenty five.
A is equal to two hundred minus one hundred and seventy five which is equal to one hundred and seventy five
Beta is equal to five hundred minus one hundred and twenty five which is equal to three hundred and seventy five
Alpha beta is equal to three hundred and seventy five minus one hundred and seventy five which is equal to two hundred and finally
Alpha is equal to hundred plus two hundred which is equal to three hundred
The coefficient of association between delinquency and income groups is given by, Q is equal to (AB) into (alpha beta) minus (A beta) into (alpha B) divided by (AB) into (alpha beta) plus (A beta) into (alpha B)
Is equal to twenty five into two hundred minus one hundred and seventy five into hundred divided by twenty five into two hundred plus one hundred and seventy five into hundred
Is equal to minus twelve thousand five hundred divided by 22 thousand five hundred
Which is equal to minus zero point five, five.
Thus there is negative association between income and delinquency.

Exercise 9
From the data given below calculate Yule's coefficient of association between weight of

children and their economic condition, and interpret it.

Figure 5

	Poor Children	Rich Children
Below Normal Weight	75	23
Above Normal weight	5	42

In the data we have given the number of children who are poor and which are how many of them are below normal weight and above normal weight.

Solution:
Let A denote poor children and B denote children below normal weight. Then alpha would denote rich children and beta children above normal weight.
The given data are (AB) is equal to seventy five, (A beta) is equal to five; (alpha B) is equal to twenty three; (alpha beta) is equal to forty two

Substituting the values in Yule's coefficient of association, we get,
Q is equal to (AB) into (alpha beta) minus (A beta) into (alpha B) divided by (AB) into (alpha beta) plus (A beta) into (alpha B)
Is equal to (seventy five) into (forty two) minus (five) into (twenty three) divided by (seventy five) into (forty two) plus (five) into (twenty three)
Is equal to three thousand thirty five divided by three thousand two hundred and sixty five which is equal to zero point nine three.

Observe that there is high degree of positive association between poor children and children below normal weight. This means that the chances of poor children being below normal weight are very high. Rich children will generally have above normal weight.

Exercise 10
Among the adult population of a certain town fifty per cent are males, sixty percent are wage earners and fifty percent are forty five years of age or over, ten percent of the males are not wage earners and forty percent of the males are under forty five. Make a best possible inference about the limits within which the percentage of persons (male or female) of forty five years or over is wage earners.

Solution:
We have been given all the figures in percentages. Hence we take N is equal to hundred.
Let us denote male by A, wage earners by B and forty five years of age or over by C.

Hence we have
N is equal to hundred, (A) is equal to fifty, (B) is equal to sixty and (C) is equal to fifty.
Further (A beta) is equal to fifty into ten percent is equal to five
(A gamma) is equal to fifty into forty percent is equal to twenty
Therefore (AB) is equal to (A) minus (A beta) is equal to forty five
(AC) is equal to (A) minus (A gamma) is equal to thirty.

We are required to find the limits for (BC).

Using the conditions of consistency,

i. (AB) plus (BC) plus (AC) is greater than or equal to (A) plus (B) plus (C) minus N
Implies (BC) is greater than or equal to (A) plus (B) plus (C) minus N minus (AB) minus (AC)
Implies (BC) is greater than or equal to fifty plus sixty plus fifty minus hundred minus forty five minus thirty is equal to minus fifteen

ii. (AB) plus (AC) minus (BC) is less than or equal to A
Implies (BC) is greater than or equal to (AB) plus (AC) minus (A) is equal to forty five plus thirty minus fifty is equal to twenty five.

iii. (AB) plus (BC) minus (AC) is less than or equal to (B)
Implies (BC) is less than or equal to (B) plus (AC) minus (AB) is equal to sixty plus thirty minus forty five is equal to forty five.

iv. (AC) plus (BC) minus (AB) is less than or equal to C
Implies (BC) is less than or equal to (C) plus (AB) minus (AC) is equal to fifty plus forty five minus thirty is equal to sixty five.

Therefore from the conditions 1 to 4, we have twenty five less than or equal to (BC) less than or equal to forty five.
Hence the percentage of wage earning population of forty five years or over must lies between twenty five and forty five.

Here's a summary of our learning in this session where we have understood how to:

- Find the missing frequencies
- Verify whether the given data is inconsistent
- Verify whether there is any association between the attributes or they are independent
- Find Yule's coefficient of association