1. Introduction and Number of Classes (Part-1)

Welcome to the series of E-learning modules on three way classified data.

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

- Explain how to classify the data when we have three attributes
- Explain the number of classes we get and their nature
- Explain how to write the classes in higher order
- Explain how to check the Consistency of the data

In the earlier modules we have discussed about the classification of attributes.

Usually attributes are classified according to possession of the attributes and non-possession of attribute by the units in the population. Hence, we always have dichotomous classification.

Possession of the attributes are denoted by capital letters A, B, C etc., and non-possession of the attributed denoted by alpha, beta, gamma etc.

In manifold classification we divide the population into more than 2 classes that is not only by possession and non-possession but also by different degree of possession.

In three way classification we consider the association of 3 characteristics or attributes.
For simplicity here also we consider the dichotomous classes. That is each attribute is divided into two classes, namely possession of attribute and non-possession of attributes.

The number of classes in three attributes can be found as follows.

As it is pointed out in earlier module that the total number of classes in which a universe can be divided on the basis of attributes is 3 to the power n.
Thus, in case of two attributes, the total numbers of classes are 3 square is 9.

Similarly, in case of three attributes the total number of classes would be 3 cube or 27. Out of these 27 classes the number of classes of the ultimate order would be 2 cube or 8.

The number of classes of the various order in case of three attributes would be as follows.

One zero order class N, which is of positive nature.

6 classes of first order of which (A), (B), (C), are of positive nature and (alpha), (beta), (gamma), are of negative nature.

12 classes of Second order of which (AB), (AC), (BC), are of positive nature. (A beta), (A gamma), (B gamma), (alpha B) (alpha C) (beta C), are pairs of contraries and (alpha beta) (alpha gamma) (beta gamma), are of negative nature.

8 classes of third order of which (ABC) of positive nature. (AB gamma), (A beta C) (alpha BC), (A beta gamma) (alpha B gamma), (alpha beta C), are pairs of contraries and (alpha beta gamma), of negative nature.

As we had seen in case of two attributes, here also, any class frequency can be expressed in terms of the frequencies of higher order. Thus,

(AB) is equal to (ABC) plus (AB gamma).
(AC) is equal to (ABC) plus (A beta C).
(BC) is equal to (ABC) plus (alpha BC) and so on.

Similarly, a higher order frequency can be expressed in terms of lower order and some higher order frequencies. Thus,

(ABC) is equal to (AB) minus (A beta C).
(A beta C) is equal to (AC) minus (ABC).
(alpha beta C) is equal to (alpha C)-(alpha BC) and so on.

Further all frequencies can be expressed in terms of the frequencies of the highest order or ultimate order. Thus,
(A) is equal to (AB) plus (A beta).
Is equal to (ABC) plus (AB gamma) plus (A beta C) plus (A beta gamma) and so on.

The following chart will clearly indicate the relationship between various classes when three attributes are involved.

2. Introduction and Number of Classes (Part-2)

From first chart, observe that N is equal to (A) plus (alpha).
Is equal to [(AB) plus (A beta)] plus [(alpha B) plus (alpha beta)].
Is equal to {[(ABC) plus (AB gamma)] plus [(A beta C) plus (A beta gamma)]} plus{[(alpha BC) plus (alpha B gamma)] plus [(alpha beta C) plus (alpha beta gamma)]}.

Similarly, we have the second chart in which B is introduced first and then C and A are introduced subsequently. Hence we can write,
N is equal to (B) plus (beta).
Is equal to [(BC) plus (B gamma)] plus [(beta C) plus (beta gamma)].
Is equal to {[(ABC) plus (alpha BC)] plus [(AB gamma) plus (alpha B gamma)]} plus {[(A beta C) plus (alpha beta C)] plus [(A beta gamma) plus (alpha beta gamma)]}.

In third chart first C is traduced and then A and B are introduced subsequently.
N is equal to (C) plus (gamma).
Is equal to [(AC) plus (alpha C)] plus [(A gamma) plus (alpha gamma)].
Is equal to {[(ABC) plus (A beta C)] plus [(alpha BC) plus (alpha beta C)]} plus {[(AB gamma) plus (A beta gamma)] plus [(alpha B gamma) plus (alpha beta gamma)]}.

From the above charts, it is clear that,

1. N is the total of all frequencies of the third order
2. (A) is the total of all the ultimate classes in which A is present. Similarly, (B) and (C) are respectively the totals of all the frequencies of the ultimate classes in which (B) and (C) are present
3. (AB), (AC) and (BC) are respectively the totals of all ultimate classes in which (AB), (AC) and (BC) are present
4. Similar relationships can be established about other classes

Now let us consider some remarks.

Let us write symbolically A into N is equal to (A), which means that the operation of dichotomizing N according to A gives the class frequency equal to (A).

Similarly, We write,
Alpha into N is equal to (α)
Adding we get,
A into N plus alpha into N is equal to (A) plus alpha
(A plus alpha) into N is equal to (A) plus (alpha) is equal to N
Implies, A plus alpha is equal to 1
Hence, we can replace A by (1 minus alpha) and alpha by (1 minus A)
Similarly with B and C.

If we know number of items possessing the given attribute, then we can fill the other columns in the contingency table using these values.
Hence, we have general relationship between the known values and unknown values as given below.

(alpha) is equal to N minus (A), (beta) is equal to N minus (B), (gamma) is equal to N minus (C).
(alpha beta) is equal to alpha beta into N.
Is equal to (1 minus A) into (1 minus B) into N.
Is equal to N minus [(A) plus (B)] plus (AB).
(alpha beta gamma) is equal to alpha into beta into gamma into N.
Is equal to (1 minus A) into (1 minus B) into (1 minus C) into N.
Is equal to N minus [(A) plus (B) plus (C)] plus [(AB) plus (AC) plus (BC)] minus (ABC).
(AB gamma) is equal to A into B into gamma into N is equal to A into B into (1 minus C) into N is equal to (AB) minus (ABC).
(A beta gamma)= A into beta into gamma into N.
Is equal to A into (1 minus B) into (1 minus C)into N.
Is equal to (A) minus (AB) minus (AC) plus (ABC).
and so on.

3. Consistency of the Data (Part-1)

Now let us test for the consistence of the data.

In case of three attributes also the given data would be inconsistent if any class-frequency is negative.
The following rules of consistence are applicable in case of three attributes.

1. (ABC) greater than or equal to zero
2. (AB gamma) is greater than or equal to zero, implies (ABC) is less than or equal to (AB)
3. (A beta C) is greater than or equal to zero implies (ABC) is less than or equal to (AC)
4. (alpha BC) is greater than or equal to zero implies (ABC) is less than or equal to (AC)
5. (A beta gamma) is greater than or equal to zero implies (ABC) is greater than or equal to (AB) plus (AC) minus (A)
6. (alpha B gamma) is greater than or equal to zero implies, (ABC) is greater than or equal to (AB) plus (BC) minus (B)
7. (alpha beta C) is greater than or equal to zero implies (ABC) is greater than or equal to (AC) plus (BC) minis (C)
8. (alpha beta gamma) is greater than or equal to zero implies (ABC) is less than or equal to (AB) plus (BC) plus (AC) minus (A) minus (B) minus(C) plus N

From the above 8 rules we can derive four additional rules by combining upper and lower limits of ABC.
Thus, (ABC) greater than or equal to zero. This is the minimum value of (ABC).
Now, (ABC) less than or equal to (AB) plus (BC) plus (AC) minus (A) minus (B) minus (C) plus N. This is the maximum value of (ABC).
The maximum value cannot be less than the minimum value and we drive a new rule,

i. (AB) plus (BC) plus (AC) is greater than or equal to (A) plus (B) plus (C) minus N.

Similarly,

ii. From 2 and 7 we get,
(AC) plus (BC) minus (AB) is less than or equal to C

iii. From 3 and 6, we get,
(AB) plus (BC) minus (AC) is less than or equal to (B) and

iv. From 4 and 5, we get,
(AB) plus (AC) minus (BC) is less than or equal to A

Hence given 8 conditions of consistency can be reduced to 4 conditions.

Consider the first Result,

One more rule of consistence applicable to any number of attribute is that for N attributes (A, B, C, D… M) is (ABC upto M) is greater than or equal to [(A) plus (B) plus (C) plus (D) upto (M)] minus (n minus 1) into N.

Let us prove this result by induction.

First we consider the two attributes A and B. Then one of the conditions for consistency is, (alpha beta) greater than or equal to zero implies
(AB) is greater than or equal to (A) plus (B) minus (N)

We prove the above condition as follows.
Consider (alpha beta) is equal to (alpha) minus (alpha B)
Is equal to N minus (A) minus [(B) minus (AB)]
Is equal to N minus (A) minus (B) plus (AB)
Implies (AB) is equal to (alpha beta) plus (A) plus (B) minus N

From this expression we can observe that, if (AB) is less than (A) plus (B) minus N, then (alpha beta) will be negative.

Hence (AB) is greater than or equal to (A) plus (B) minus N

Applying above inequality to the universe of C we get,
(ABC) is greater than or equal to (AC) plus (BC) minus (C).
Greater than or equal to (A) plus (C) minus N plus (B) plus (C) minus N minus (C).
Greater than or equal to (A) plus (B) plus (C) minus 2 into N.

Applying the above inequality to the universe of D, we get,
(ABC) is greater than or equal to (AD) plus (BD) plus (CD) minus 2 into D.
Greater than or equal to (A) plus (B) minus N plus (B) plus (D) minus N plus (C) plus (D) minus N minus 2 into D.
Greater than or equal to (A) plus (B) plus (C) plus (D) minus 3 into N.

Therefore, in general we can write for n attributes as,
(ABCD upto M) is greater than or equal to (A) plus (B) plus (C) plus (D) upto M) minus (n minus 1) into N.

4. Consistency of the Data (Part-2)

Consider the 2nd result.
Show that,

i. If all A's are B' and all B's are C's, then all A's are C's
ii. If all A's are B's and no B's are C's then no A's are C's

Among two results we prove first one as follows.

We have given all A's are B' implies (AB) is equal to (A)
and all B's are C's implies (BC) is equal to (B)

To prove (AC) is equal to (A),
We have, (AB) plus (BC) minus (AC) is less than or equal to (B).
Implies (A) plus (B) minus (AC) is less than or equal to (B)
Implies (A) is less than or equal to (AC) or (AC) is greater than or equal to (A)

But since (AC) cannot be greater than (A), we get (AC) is equal to A.

In the second, we have given,

All A's are B's implies (AB) is equal to (A).
And no B's are C's implies (BC) is equal to zero and we want to prove that (AC) is equal to zero.

We have (AB) plus (AC) minus (BC) is less than or equal to A.
implies (A) plus (AC) minus zero is less than or equal to (A).
Implies (AC) is less than or equal to zero.

And since (AC) is greater than or equal to zero, we must have (AC) is equal to zero.

Now consider the third result.

Show that, if
(A) by N is equal to x, (B) by N is equal to 2 into x and (C) by N is equal to 3 into x,
And (AB) by N is equal to (BC) by N is equal to (CA) by N is equal to y,
Then the value of neither x nor y can exceed 1 by 4.

To prove this we consider the conditions of consistency.
That is (AB) is less than or equal to (A) implies N into y is less than or equal to N into x implies y is less than or equal to x. Name this as 1.

Also (BC) is greater than or equal to (B) plus (C) minus N.
If we divide both the sides by N, we get,
(BC) by N is greater than equal to (B) by N plus (C) by N minus 1.

Implies y is greater than or equal to 2 into x plus 3 into x minus 1.

Therefore, 5 into x minus 1 is less than or equal to y. Name it as 2.
1 and 2 gives, 5 into x minus 1 is less than or equal to x.
Implies 4 into x is less than or equal to 1.
Implies x is less than or equal to 1 by 4. Name it as 3.

Thus from 1, 2 and 3, we have y is less than or equal to x which is less than or equal to 1 by 4, which establishes the result.

Now let us consider the fourth result.

Given that (A) is equal to (alpha) is equal to (B) is equal to (beta) is equal to (C) is equal to (gamma) is equal to half into N and also (ABC) is equal to (alpha beta gamma), show that, 2 into (ABC) is equal to (AB) plus (AC) plus (BC) minus half into N.

The above result is proved as follows.

From the remark we know that,
(alpha beta gamma) is equal to N minus [(A) plus (B) plus (C)] plus [(AB) plus (AC) plus (BC)] minus (ABC).

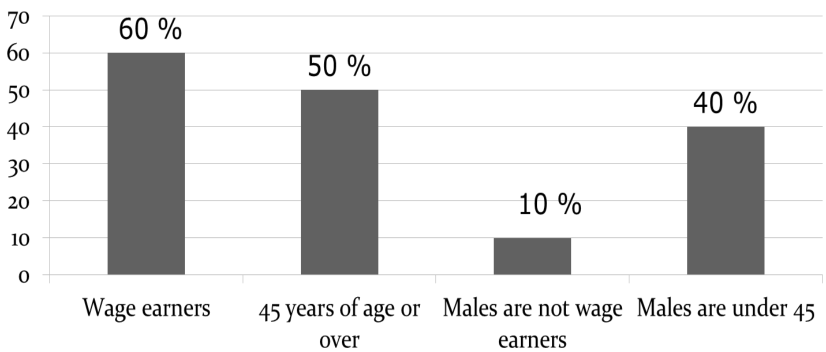
Substituting for (A), (B), (C) and (alpha beta gamma), we get,
(ABC)=N minus [half into N plus half into N plus half into N] plus [(AB) plus (AC) plus (BC)] minus (ABC).
Implies, 2 into (ABC) is equal to N minus (3 into N by 2) plus (AB) plus (AC) plus (BC)
Implies, 2 into (ABC) is equal to (AB) plus (AC) plus (BC) minus half into N

5. Illustration

Now let us consider one illustration.

Among the adult population of a certain town 50 per cent are males, 60 percent are wage earners and 50 percent are 45 years of age or over, 10 percent of the males are not wage earners and 40 percent of the males are under 45. Make a best possible inference about the limits within which the percentage of persons (male or female) of 45 years or over is wage earners.

Figure 1



We have given all the figures in percentages. Hence, we take N is equal to 100.

Let us denote male by A, wage earners by B and 45 years of age or over by C.

Hence we have given,
N is equal to 100, (A) is equal to 50, (B) is equal to 60 and (C) is equal to 50.

Further, (A beta) is equal to 50 into 10 percent of 50 is equal to 5.
(A gamma) is equal to 50 into 40 percent of 50 is equal to 20.

Therefore, (AB) is equal to (A) minus (A beta) is equal to 45.
(AC) is equal to (A) minus (A gamma) is equal to 30.

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 50 plus 60 plus 50 minus 100 minus 45 minus 30 is equal to minus 15.
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 45 plus 30

minus 50 is equal to 25.

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 60 plus 30 minus 45 is equal to 45.

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 50 plus 45 minus 30 is equal to 65.

Therefore from the conditions 1 to 4, we have, 25 less than or equal to (BC) less than or equal to 45.

Hence, the percentage of wage earning population of 45 years or over must lies between 25 and 45.

Here's a summary of our learning in this session, where we understood:

- How to classify the data when we have three attributes
- The Number of classes we get and their nature
- How to write the classes in higher order
- How to check the Consistency of the data