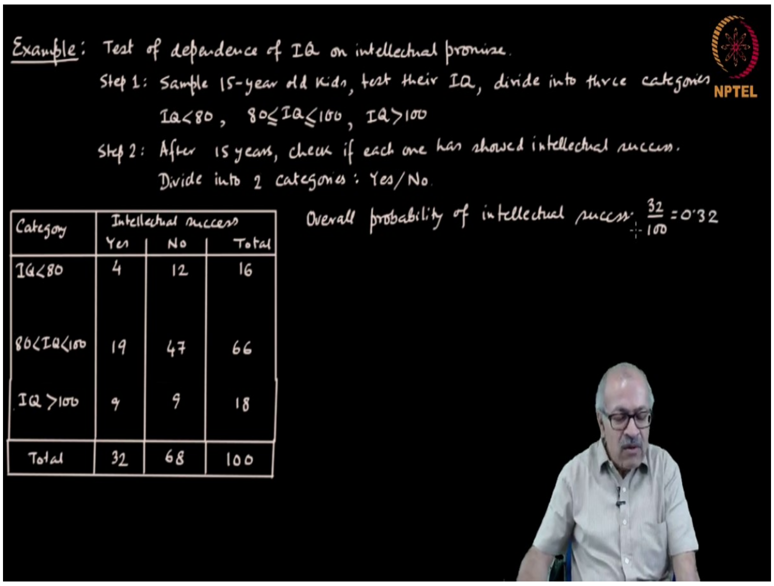
Research Methodology
Prof. Soumitro Banerjee
Department of Physical Sciences
Indian Institute of Science Education and Research, Kolkata

Lecture - 47
Hypothesis Testing: The Chi-Square Test Part 03

(Refer Slide Time: 00:17)



In the last class, we learned how to apply the chi square technique of hypothesis testing and we saw an example. Today let us continue with that with another example.

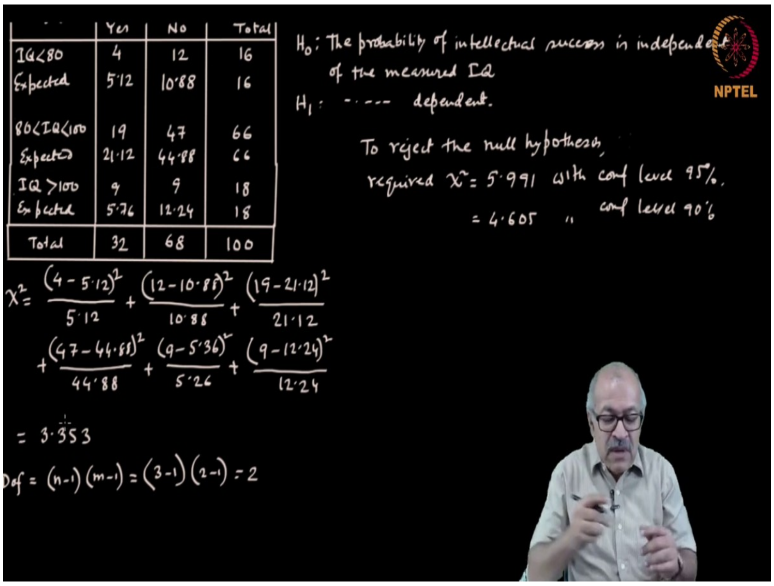
It is normally believed that those who have a high IQ measured at a relatively young age will have a larger promise of intellectual success at a later age. Suppose you are setting out to test that belief. How would you do that? This is how it is conducted. Suppose that, first you get a collection of 15-year-old kids from a school, more or less of the same academic background as well as financial background, so that their financial status, social status, is more or less the same.

And suppose you have got some 100 kids like that, you test their IQ and divide them into three categories; IQ less than 80, IQ between 80 and 100, and IQ greater than 100. Then you wait for 15 years. After 15 years, you check the intellectual success of each one of these 100 kids, and divide them into two categories depending on the answer to the question: “did they excel in any intellectual pursuit?” Yes or no? Accordingly you divide.

Suppose you got the result something like this. Initially when you made the test, 16 kids were in the IQ less than 80 category and out of that 4 have excelled in intellectual pursuit, 12 have not. In the middle one, that means from 80 to 100 IQ, there are 66 kids; out of whom 19 excelled in some kind of intellectual pursuit and 47 did not. And 18 kids had IQ greater than 100 and out of that 9 excelled and 9 did not.

On that basis, can you infer that IQ is an indicator of promise of intellectual success? This is the problem that we are tackling. Now, you notice that out of a 100 kids, 32 excelled in intellectual promise. So, overall probability of intellectual success is 32 by 100 is equal to 0.32.

(Refer Slide Time: 03:36)



Now, in this case what is the null hypothesis and what is the alternative hypothesis? The null hypothesis H0 is that the probability of intellectual success is independent of the measured IQ. And what is the alternative hypothesis? That says ‘dependent’, so the probability of intellectual success is dependent of the measured IQ.

Now, we have said that whenever we set out to test a hypothesis, we always root ourselves on the null hypothesis and check if there is enough evidence to reject the null hypothesis. So, if the null hypothesis is true, then our expectation would be that in all these categories, there would be the same ratio of intellectually successful people. So, this ratio will remain the same irrespective of whether one is in this category, this category or this category. From there we can draw the expectations. Let me write the expected numbers.

So, out of 16, 0.32 times 16 should be expected to find intellectual success. That comes to be 5.12. The expected number in the 'no' category will be 16 minus 5.12 = 10.88. So, the remaining number should not be expected to be intellectually successful and these two add up to 16.

Similarly here the probability of intellectual success of all 66 people should be the same if the null hypothesis true. So, 66 times 0.32 is 21.12. And 66 minus 21.12, that many number will not be expected to be successful, that number is 44.88. This total is also 66.

Similarly, there also will be one expected; that number will be 18 times 0.32, which is 5.76. And 18 minus 5.76, would be 12.24. And you would notice that all these rows add up to these numbers, and all these columns also add up to this number.

Categories	Intellectual success		
	Yes	No	Total
IQ<80	4	12	16
Expected	5.12	10.88	16
80≤IQ≤100	19	47	66
Expected	21.12	44.88	66
IQ>100	9	9	18
Expected	5.76	12.24	18
Total	32	68	100

So, 4 plus 19 plus 9 is equal to 32, 12 plus 47 plus 9 is equal to 68 and so on and so forth. So, this is a contingency table for this particular problem. Now we set out to do the chi square test. Now, how would you do the chi square test? We have to write the value of chi square.

So, let me put it like this. The chi square, as you know, is a sum of the expected minus the observed square, divided by the expected. So, the first row will be 4 minus 5.12 whole square divided by 5.12. Plus 12 minus 10.88 square by 10.88. Now, the second row: 19 minus 21.12 square by 21.12 plus 47 minus 44.88 square by 44.88 plus 9 minus 5.36 by 5.36; this one will be 9 minus 12.24 square by 12.24. This calculates to 3.353.

$$\begin{aligned}\chi^2 &= \frac{(4-5.12)^2}{5.12}+\frac{(12-10.88)^2}{10.88}+\frac{(19-21.12)^2}{21.12}\\ &\quad+\frac{(47-44.88)^2}{44.88}+\frac{(9-5.76)^2}{5.76}+\frac{(9-12.24)^2}{12.24}\\ &= 3.353\end{aligned}$$

So, now this is the chi square value for this particular contingency table. If this is a chi square value, then we will consult the chi square table to find out if we can reject the null hypothesis. But before we do that, we need to find out the degree of freedom. What is the degree of freedom in this case? There are three rows; but then if you know the total number, by knowing two of them you can calculate the third. So, the third one does not have a degree of freedom. So, vertically, around the rows, there are 2 degrees of freedom.

Along the columns: there are two columns, but if you know the total and if you know one, you know the other. Therefore, along the vertical you have only 1 degree of freedom. We said it that it is basically n minus 1 times m minus 1; in this case it is 3 minus 1 times 2 minus 1 is equal to 2. So, it is only 2 degrees of freedom.

So, now with that information, we consult the chi square table. I have already shown you in the last class how to consult the chi square table. You consult the chi square table and find that, in order to reject the null hypothesis with 95 percent confidence, the required chi square will be 5.991. That means significance level 0.05.

You notice that the value 3.353 is below that. If the confidence level is even lowered to 90 percent, then the required value of chi square would be 4.605. It is even below that, which means that the differences, even though to a untrained eye it appears that among the number that is below 80 IQ, a larger fraction did not perform intellectually, but these variations appear to be satisfying the requirements of a random distribution. Here we cannot reject the null hypothesis with 95 percent confidence, we cannot even with 90 percent confidence. So we cannot reject the null hypothesis.

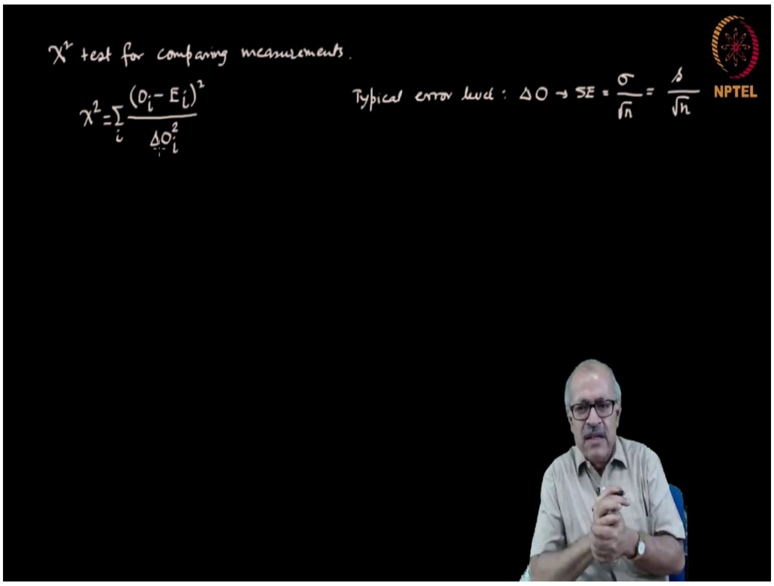
And so, on the basis of this data we would say that, we do not have enough evidence to reject the null hypothesis, and on that basis we will accept the null hypothesis. There are factors other than IQ that are principally responsible for one's intellectual success in life: the way an individual conducts a life, how he learns from the surroundings even after the age of 50. So, these are, in many cases, the more important determining factors in deciding the intellectual success in one's life.

Now, so far we were dealing with the chi square tests where we were putting individuals in categories: the tall plants, the short plants are the categories, an IQ level below 80 is a bin, a category, in which you can put an individual. So, in that kind of situations, we have learnt how to do the test.

But there are also situations, especially in natural sciences, where a hypothesis or a theory predicts a value; not whether how many can be put in a bin, but rather a value. How do you test that kind of a hypothesis or a theory in that situation, where it is a value that is predicted?

Firstly, our logic was that, we will subtract the expected value from the observed value. We will perform the test experiment and then, for that particular experimental condition there should be some expectation from the theory. The expected value, and the observed value, the subtraction of that. It does not matter whether you subtract this from that or that from this, because ultimately you take a square, so that it becomes a positive number.

(Refer Slide Time: 15:38)



So, here we are talking about chi square test, for comparing measurements with some expected value. There is some expected value. If an experiment has an expected value E and the observed value O, then we are talking about the observed value minus expected value, squared.

Now, we have learnt that this needs to be normalized. In this case what will you normalize it with? Now, the difference between the observed value and the expected value could be due to some kind of a systematic factor, for example, the theory being wrong, or it could be due to random errors. And therefore, it is logical to compare that with the typical error level of that particular experiment.

The typical error level in an experiment is an error in observation. Now, what is that? We have learnt earlier that if you make a measurements a number of times and take their mean, then from there you can find out the standard deviation and the standard deviation divided by square root n is the standard error of the mean. The standard error of the mean is a typical indicator of the error level of the experiment. So, this is essentially the standard error of the mean. We know how to obtain that.

So, we will have to normalize it by the standard error, delta O. But in the case where we are making measurements, these could have units, for example, centimetres, grams per second or whatever. So, you have these in units; but the chi square should have no units, it should be non-dimensional. Therefore, you have to divide it by delta O square. So, that would be the chi square.

Now, this is true if you are able to make one measurement. But a theory may predict different outcomes for different conditions in a measurement and therefore, we have to make provisions for a number of such measurements being done and for each one there is an expected value. And if we keep that provision, it would have a subscript i, the ith measurement, and the ith prediction, and then you have to sum over all i's.

Suppose a theory predicts m number of different situations, and for each situation, there is an expected value. You measure it and you get the corresponding observed value and you also get the corresponding standard error of the mean.

And then this will represent, in that case, the chi square value, and the rest of the test is the same. Remember that in that case the standard error has to be measured in a proper way as we have learnt earlier; which means you have to take 25 readings, otherwise it will not be distributed as a normal distribution.

The normal distribution has to be guaranteed in order for this to work and then you extract the standard error of the mean, which is basically the sigma by square root of n. And if you do not know the sigma, then it is the measured standard deviation divided by square root of n. So, this will have to be substituted here. Now, let us illustrate this method using an example.