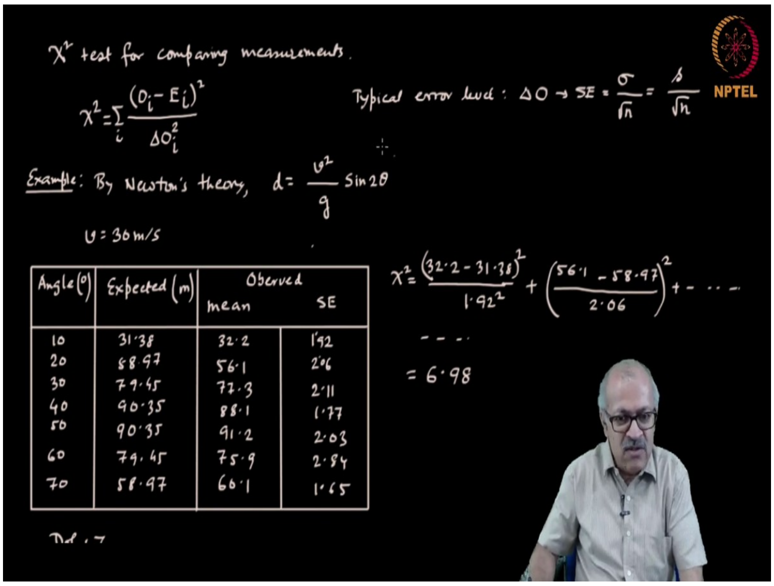
Research Methodology
Prof. Soumitro Banerjee
Department of Physical Sciences
Indian Institute of Science Education and Research, Kolkata

Lecture - 48
Hypothesis Testing: The Chi-Square Test, Part 04

(Refer Slide Time: 00:18)



Now, let us do an example. You know the Newton's theory makes predictions regarding the how far will a cannonball go for different angles of the cannon. If it is a projectile thrown with a velocity v at an angle theta, then the distance to which it will go by Newton's theory, say d, is equal to v square divided by g into sin 2 theta.

$$d = \frac{v^2}{g} \sin 2\theta$$

You are trying to test whether this theory is correct or not. What will you do? You will simply take a gun and you will place it at different angles, say, 10 degrees, 20 degrees, 30 degrees, 40 degrees, so on and so forth. For each one you will fire it and depending on the character of the gun it will fire at a particular speed. Suppose that the speed v is 30 meters per second. That is constant because, you are taking a single gun to do that.

In that case, what will the contingency table look like? The contingency table will look like this. First the angle and then there are three things we have to record. The observed one and what is expected. First let us write what are expected. This angle is in degrees, and this is in meters. Then there would be two things coming from observation. You have to write the mean and you have to write the standard error of the mean. I draw a line and then write the results.

Angle (degrees)	Expected (m)	Observed	
		Mean	SE
10	31.38	32.2	1.92
20	58.97	56.1	2.06
30	79.45	77.3	2.11
40	90.35	88.1	1.77
50	90.35	91.2	2.03
60	79.45	75.9	2.84
70	58.97	60.1	1.65

Obviously, there would be 10, 20, 30, 40, 50, 60, 70; suppose you have taken readings for these. There is no point pointing vertically, at 90 degrees. So, suppose you have taken this many readings. And the expected values can be calculated because v is this, g value you know, theta values are these. So, you can easily calculate and these are the expected values; 31.38, 58.97, 79.45, 90.35 ....

You know that at 45 degrees it will go to the maximum distance, and 40 and 50 are equally distant from that. Therefore, these must be equal; 90.35, 90.35 and then 79.45 and this is 58.97. From whatever you observed, the mean comes out to be 32.2 and is is 1.92, this is 56.1, this is 2.06, this is 77.3, this is 2.11, this is 88.1, this is 1.77 standard error, this is 91.2, 2.03, this is 75.9, this is 2.84, this is 60.1, this is 1.65.

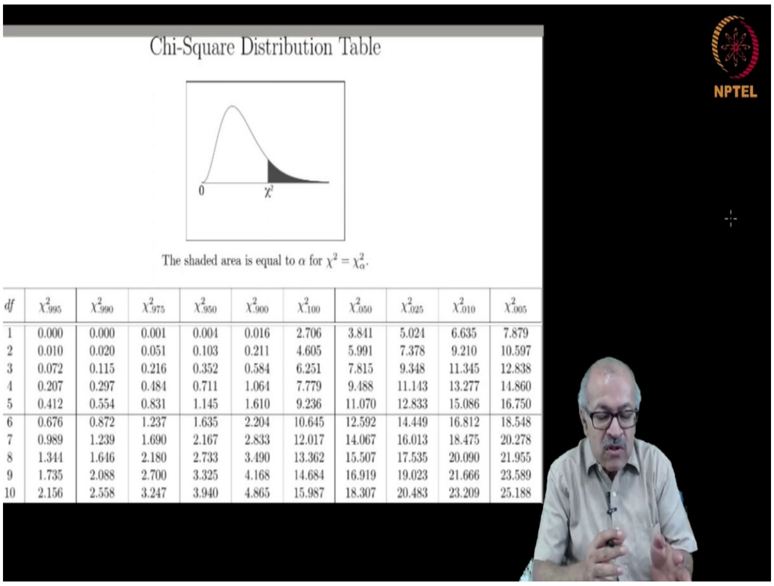
You see that the results actually differed from what you expected, they were not exactly the same, but then our test should tell us whether we can expect those differences out of random errors that can happen. So, we now write down the result based on this contingency table and here we will have to write chi-square. The chi-square would be the observed minus the expected divide by the standard error.

So, it is 32.2 minus 31.38 square by this is 1.92 square plus the second one 56.1 minus 58.97 divided by 2.06 you can put the whole one square because the numerator as well as denominator have squares. And I will not write to all of them. You can easily fill up the rest and if you do that you will find that there will be 7 of these kind of terms, it is easy to write all of them and then if you do that you get 6.98 as a total value.

$$
\begin{aligned}
\chi^2 &= \left(\frac{32.2-31.38}{1.92}\right)^2 + \left(\frac{56.1-58.97}{2.06}\right)^2 + \left(\frac{77.3-79.45}{2.11}\right)^2 \\
&\quad + \left(\frac{88.1-90.35}{1.77}\right)^2 + \left(\frac{91.2-90.35}{2.03}\right)^2 + \left(\frac{75.9-79.45}{2.84}\right)^2 \\
&\quad + \left(\frac{60.1-58.97}{1.65}\right)^2 \\
&\approx 6.98
\end{aligned}
$$

Now, we have to consult the chi-square table. What is the degree of freedom in this case? Now, notice that there are 7 rows, but then if you measure 6 of them you can say nothing about the 7th one and therefore, all 7 are degrees of freedom. So, the degree of freedom is actually 7 in this case, which was not happening for the category cases, but in this case, where you are measuring something, then all of them are degrees of freedom.

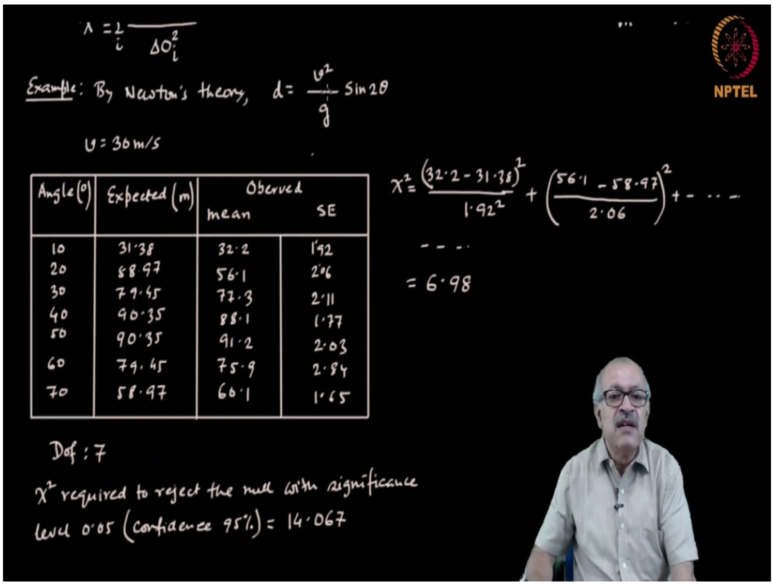
(Refer Slide Time: 08:41)



df	$\chi^2_{.995}$	$\chi^2_{.990}$	$\chi^2_{.975}$	$\chi^2_{.950}$	$\chi^2_{.900}$	$\chi^2_{.100}$	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188

So, for 7 degrees of freedom, if you now consult the chi-square distribution table, then you would notice that we have to look at the degrees of freedom 7. And since we are trying to obtain some conclusion with 95 percent confidence, which is 0.05 significance level, so we go to this 0.05 significance level and we come down and the intersection between this and that happens at 14.067. So, on that basis we have to make our decision.

(Refer Slide Time: 09:27)

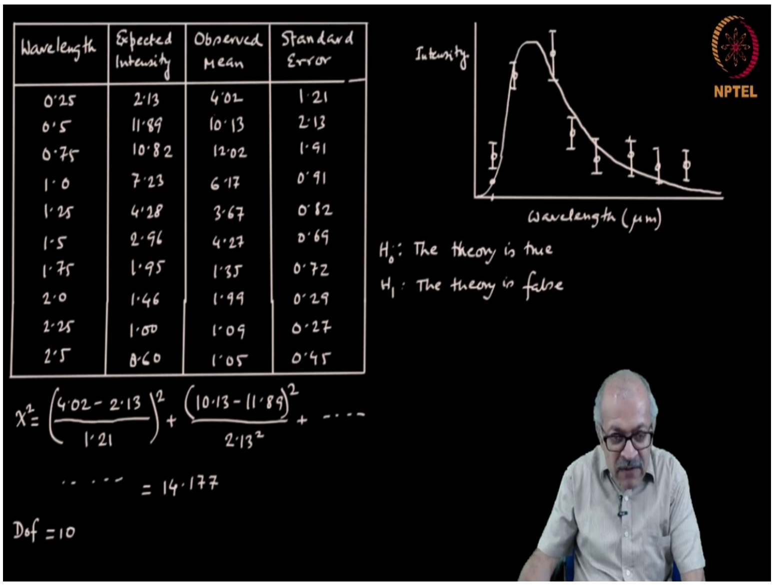


Then the chi-square value required to reject the null hypothesis is 14.067. By the way, I did not write the null and the alternative hypotheses. The null hypothesis is always an equality hypothesis. Now, here the theory predicts something and therefore, the null hypothesis says that the theory is true, and the alternative hypotheses will say the theory is false, because we have to always talk about a equality criterion and this gives an equality criterion.

So, the null hypothesis would be that the theory is true. That means, it is actually given by this. So, the chi-square value required to reject the null hypothesis with significance level 0.05, that is confidence level 95 percent, from the chi-square table that turns out to be 14.067. The value of chi square 6.98 is far smaller than that.

So, we conclude that even though the values that we actually obtained appear to be different, but these differences are due to random errors; they satisfy the behaviour of random fluctuations. So, we conclude that the null hypothesis cannot be rejected, and the Newton's theory, which predicts these values of the distance traversed, is correct.

(Refer Slide Time: 11:37)



Now, in order to understand this method a little more elaborately, let us take another example. You know that, if there is a black body at a certain temperature, it radiates energy. The energy is not radiated in a single wavelength; it is distributed over a large number of wavelengths. And towards the end of the 19th century, this distribution of intensity at different energy levels was measured.

For a long time there was no proper explanation and finally, Planck postulated that light is emitted in packets of energy. By assuming that, he derived an expression for a curve and the curve looks something like this. Now suppose after that somebody, some scientist, wants to check whether the curve is correct or not, the theory is correct or not. For that purpose, he or she again re-does the experiment.

Therefore, there is a hot body at a certain temperature, and then she measures the intensity of radiation coming at different wavelengths. At a certain wavelength it is this much, again that measurement is done again and again, that means, at least 25 times, so that she gets a mean, she gets a standard error of the mean. So, she gets a value with an error bar.

Again at a different value of the wavelength, she gets a mean with an error bar, and so on and so forth. She gets data that are distributed like this. And again in this case it is clear that the data points, the means, are not really on the graph. Then would you say that the experiment does not satisfy the theory? That is what we need to check.

Suppose the data that are expected from the theory for these wavelengths (in micrometers): what are expected are as follows. This is 2.13, this is 11.89, this is steep rise this side, and relatively a slower fall the other side; so, 10.82, 7.23, 4.28, 2.96, 1.95, 1.46, 1.00, and 0.60.

Wavelength	Expected intensity	Observed mean	Standard error
0.25	2.13	4.02	1.21
0.5	11.89	10.13	2.13
0.75	10.82	12.02	1.91
1.0	7.23	6.17	0.91
1.25	4.28	3.67	0.82
1.5	2.96	4.27	0.69
1.75	1.95	1.35	0.72
2.0	1.46	1.99	0.29
2.25	1.00	1.09	0.27
2.5	0.60	1.05	0.45

So, on the graph at different values: for example, at this value, this is the value that is expected from the graph, but this is the value that was obtained. So, this value is what I have tabulated here.

The observed means are the means that are actually experimentally obtained and this is 4.02, almost twice. These are 10.13, 12.02, 6.17, 3.67, 4.27, 1.35, 1.99, 1.09 and 1.05. So, you see in some cases there is significant difference between what is expected and what is actually observed. And the standard errors that she obtained are as follows; 1.21, 2.13, 1.91, 0.91, 0.82, 0.69, 0.72, 0.29, 0.27, and 0.45. These are the results that are obtained.

Now we have the task of checking whether these differences that she got can come from some random distributed error. And if that is so, we can say that these errors are random errors and not a kind of systematic error which might happen if the theory is wrong. So, again in this case, in order to test the theory, we will have to write the chi-square value.

So, here the chi-square will be, again the same way, 4.02 minus 2.13 divided by 1.21 here whole square plus the next one 10.13 minus 11.89... It does not really matter whether you write this one first or this one first because the ultimately you are squaring it, this is 2.13 square. I mean you would could have put a whole square. All these you have to write down.

$$
\begin{aligned}
\chi^2 \quad = \quad & \frac{(4.02-2.13)^2}{1.21^2}+\frac{(10.13-11.89)^2}{2.13^2}+\frac{(12.02-10.82)^2}{1.91^2} \\
& +\frac{(6.17-7.23)^2}{0.91^2}+\frac{(3.67-4.28)^2}{0.82^2}+\frac{(4.27-2.96)^2}{0.69^2} \\
& +\frac{(1.35-1.95)^2}{0.72^2}+\frac{(1.99-1.46)^2}{0.29^2}+\frac{(1.09-1.00)^2}{0.27^2} \\
& +\frac{(1.05-0.60)^2}{0.45^2} \\
= \quad & 14.177
\end{aligned}
$$

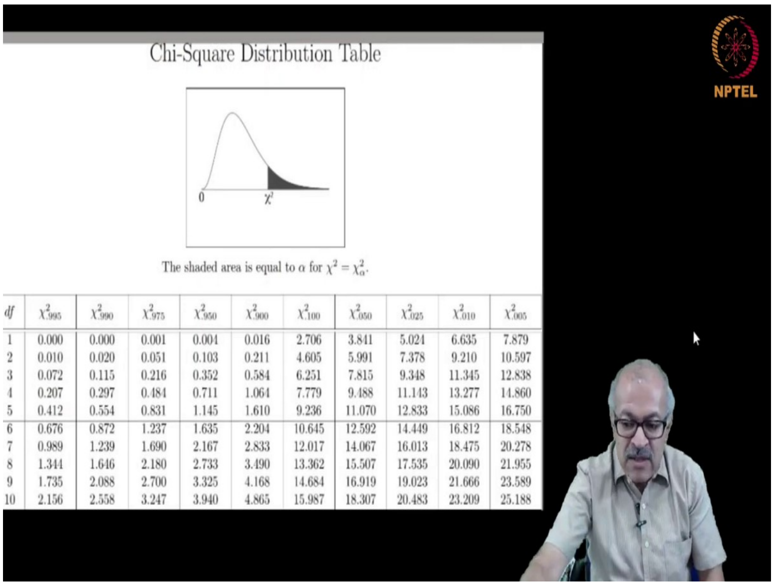
How many terms will there be? 1, 2, 3, 4, 5, 6, 7, 8, 9, 10; there will be 10 such additive terms and if you write down all of them, finally you will get 14.177.

Now, you now have to consult the chi-square table with this value with the degree of freedom. The degree of freedom will be, in this case 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 and by knowing 9 of them you do not know anything about the 10th one. Therefore, all 10 have degrees of freedom.

So, k is equal to 10, the degree of freedom is 10. For that, we will have to check from the chi-square table whether the null hypothesis can be rejected. Again the null hypothesis, in this case, will be that the theory is true and the alternative hypothesis is that the theory is false.

And you have started by believing the null hypothesis, because that is how you built up these expected intensities. That comes from the assumption of the null hypothesis and that is what we are ultimately testing: whether that is true or not.

(Refer Slide Time: 20:30)



df	$\chi^2_{.995}$	$\chi^2_{.990}$	$\chi^2_{.975}$	$\chi^2_{.950}$	$\chi^2_{.900}$	$\chi^2_{.100}$	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188

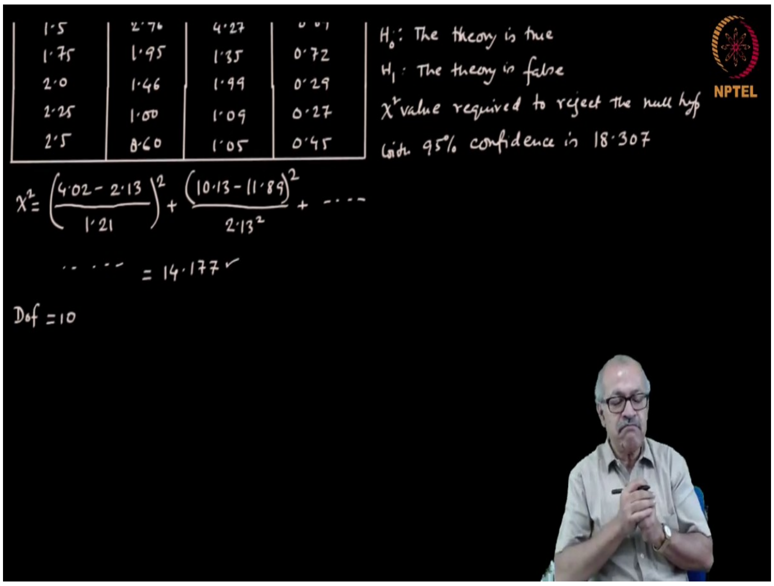
Let us now look at the chi-square table as I have shown here. The chi-square distribution would normally be like this for relatively larger values of the degree of freedom. The chi-square value for which 0.05 is the level of significance, that is related to 95 percent confidence level.

So, we will have to look at this and the value of chi-square for which the area to the right of this point is 0.05 of the whole area. That is tabulated here. So, in this particular case we have the degree of freedom as 10; so, it is here and we come to the right and to the column of 0.05 degree of significance and we find that 18.307 is the required value of chi-square to reject the null hypothesis.

So, if the chi-square value in an experiment is beyond that, we have sufficient reason to reject the null hypothesis. But, if it is on this side, that is, smaller than the critical chi-square value required for rejection of the hypothesis, that means, that the deviation between the expected value and the observed value these are due to random errors.

The differences follow the characteristic properties of a random variable; random variables that are distributed as a normal distribution with mean 0 and standard deviation 1; it more or less satisfies the character. If we get a value of chi-square which is very unexpected, then we would say that we have sufficient reason to reject the null hypothesis.

(Refer Slide Time: 23:08)



But, in this case the chi-square value required to reject the null hypothesis with 95 percent confidence is 18.307. But, the value that we actually got is 14.177 which is smaller than that.

So, we cannot reject the null hypothesis and we have to say that the data could not negate or falsify the theory that has been developed. So, that is what we have to claim at the end of the day. This is how we do the chi-square test for hypothesis testing.

It requires a bit of practice in order to understand and apply it in real life situations. So, I would request all the students to solve problems and even to imagine problems and imaginary data and check whether, on the basis of that data, the hypothesis can be rejected or not.