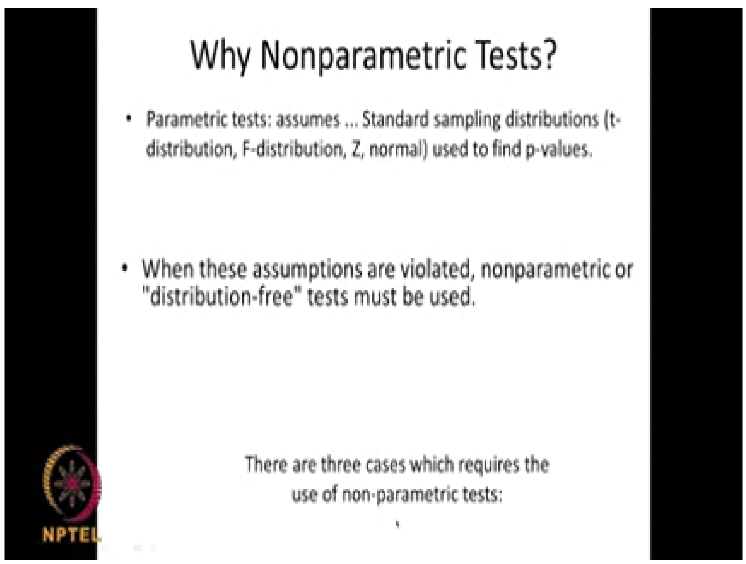
Biostatistics and Design of Experiments
Prof. Mukesh Doble
Department of Biotechnology
Indian Institute of Technology, Madras

Lecture - 28
Nonparametric Tests

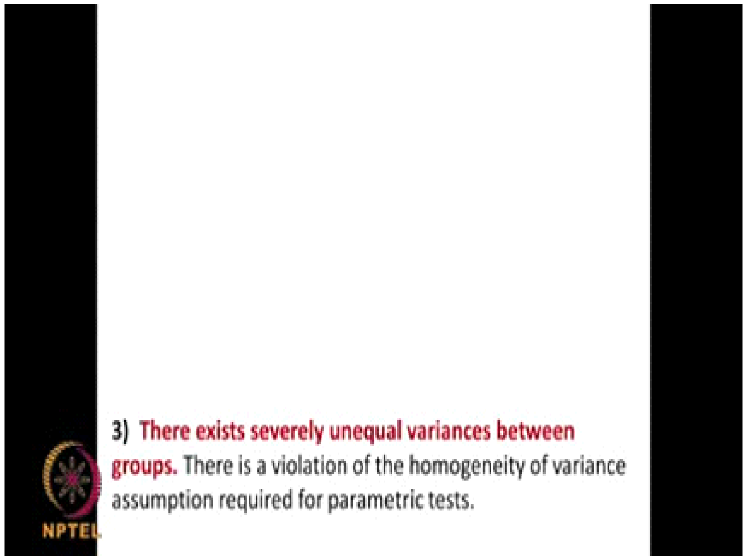
Hello everyone, welcome to the course on Biostatistics and Design of Experiments. Today, I am going to talk about Nonparametric tests. So far, we looked at parametric tests, but today it is going to be Nonparametric. What is this Nonparametric test?

(Refer Slide Time: 00:28)



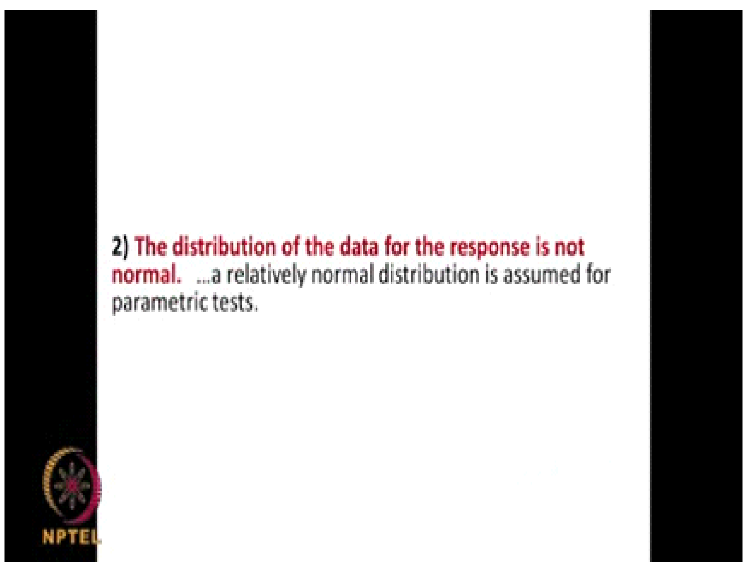
So, it assumes, in a parameter standard sampling distribution, like your t distribution, F-distribution, Z distribution, normal distribution, all these are called a standard parametric test and you have corresponding tables. And then, we used those tables to find out the p values, given the degrees of freedom and so on. But then, you can always have data which does not follow this type of distribution, actually. They are distribution free data. We will look at them, what are these distribution free data. Then, we still need to use some sort of a test for comparing distribution free data, right?. So, they are called Nonparametric test. So, there are 3 cases which require the use of nonparametric tests. What are those 3 cases?

(Refer Slide Time: 01:15)



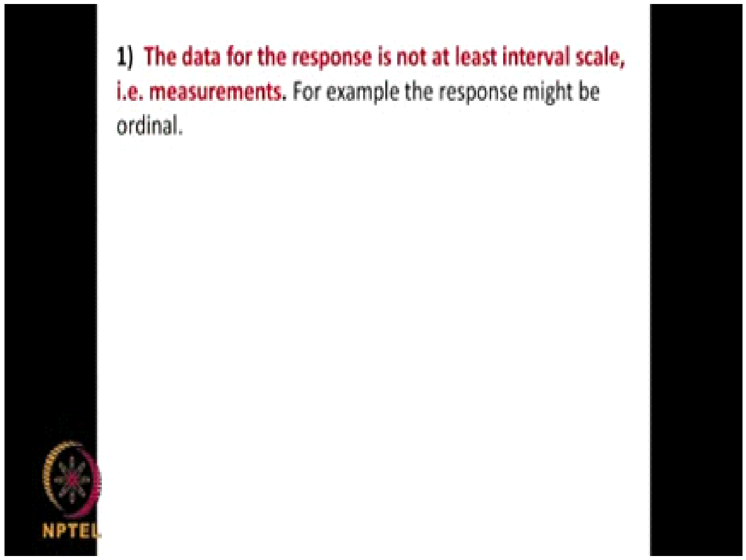
There exists severely unequal variances between the groups; that means, they do not follow the homogeneity of variance. What is that homogeneity of variance? That means, we assume that the variance is almost similar from each group.

(Refer Slide Time: 01:29)



The distribution of the data for the response is not normal. That means, they do not follow a normal distribution. Normal is mean, median, mode, are all the same, but here, they are very different, ok?.

(Refer Slide Time: 01:43)



Or, they have not, it is basically ordinal; the data could be ordinal; that means, it is not interval scale.There are no interval scale type of measurements, like we have pH as a function of time, temperature as a function of time. So, you have the time as your interval. In such situations, then obviously, we need to use something else, and they are called Nonparametric type of data, and so we need to use Nonparametric test.

(Refer Slide Time: 02:04)

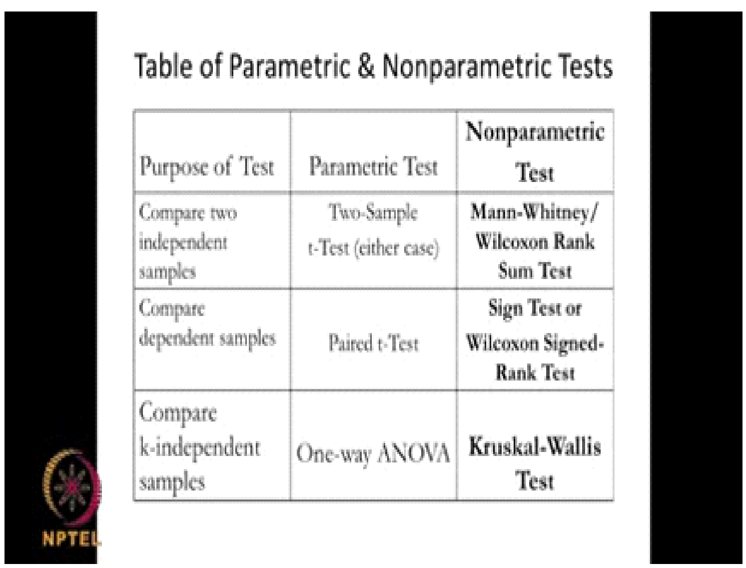


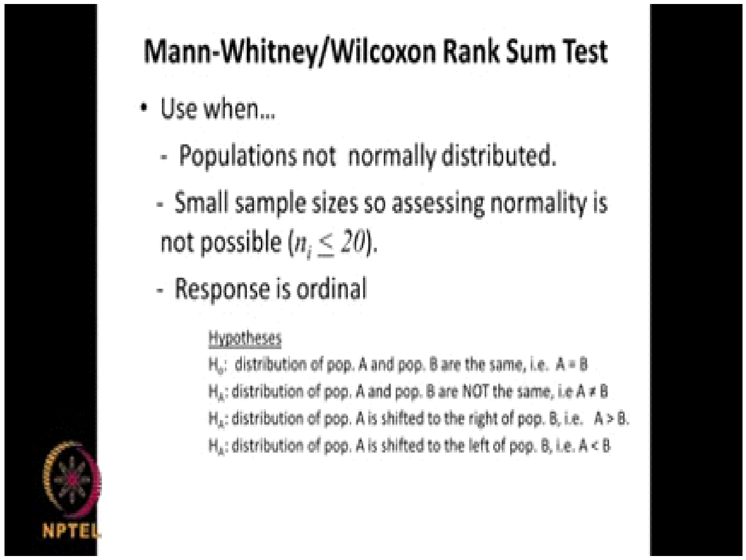
Table of Parametric & Nonparametric Tests

Purpose of Test	Parametric Test	Nonparametric Test
Compare two independent samples	Two-Sample t-Test (either case)	**Mann-Whitney/ Wilcoxon Rank Sum Test**
Compare dependent samples	Paired t-Test	**Sign Test or Wilcoxon Signed-Rank Test**
Compare k-independent samples	One-way ANOVA	**Kruskal-Wallis Test**

So, homogeneity of variance is violated, and the data is not normal, or the data is ordinal, that means, there is no x axis like a time, for example; you know, pH versus time, or temperature versus time. It is more of an ordinal time. Then, we use the nonparametric type of test. You have, similar to each of the parametric test a corresponding nonparametric test and we will spend some time on each one of them as well. So suppose, I am comparing two independent samples, generally, we talked quite a lot, right?, Two sample t-test; then, for a nonparametric, we use something called Mann-Whitney or Wilcoxon Rank Sum Test, Wilcoxon Rank Sum Test.

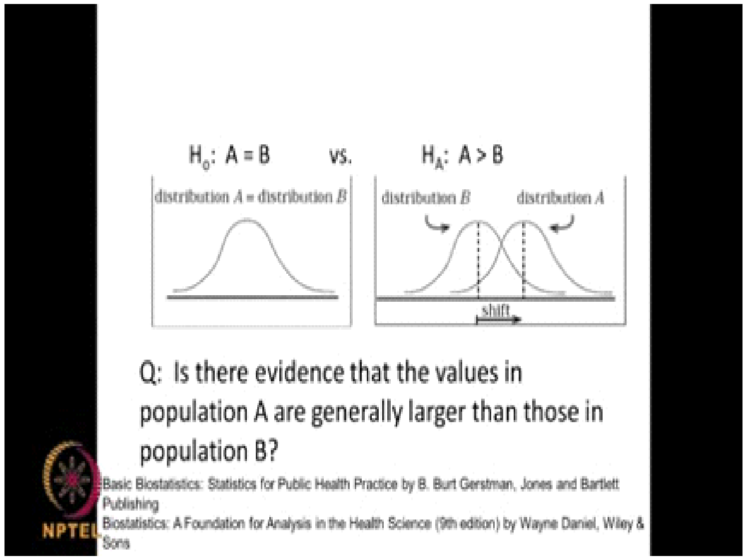
If you are comparing dependent samples, that means, the Paired t-test, you are using the same subjects for giving the placebo as well as the drug, you are the using same subjects for giving the drug a and drug b, then, we use corresponding nonparametric test is Sign test or Wilcoxon Signed Rank Test. If you are comparing k-independent samples, that means, you are comparing many drugs, I want to see whether there is a statistical difference between drugs, we used to do ANOVA, One-way ANOVA, right?; you remember all this very well. Corresponding nonparametric test is Kruskal-Wallis Test. So, these are the tests you need to know, especially for nonparametric data. Like I said, when do you know it is nonparametric? If the homogeneity of variance is violated, or it is not like a normal distribution, or some sort of a distribution like t, or f, or Z type of distribution, and then, the data is ordinal, there is no variation with respect to time or something like that, then we need to use these type of tests to determine, for comparing independent samples, comparing dependent samples, looking at large number of samples, and so on, actually. Later on, we will also talk about what is this homogeneity of variance, how do you find out, later on. But as of now, we will talk about these various Nonparameter based test.

(Refer Slide Time: 04:26)



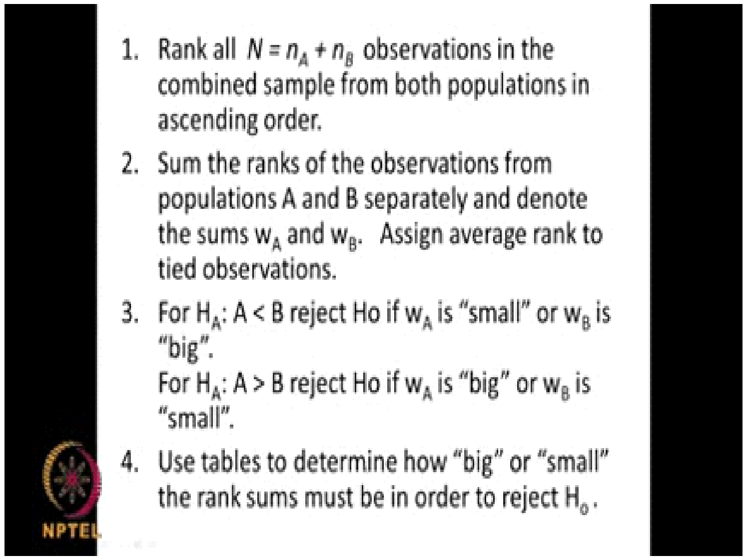
So, let us look at this Mann-Whitney/Wilcoxon Rank Sum Test. When do you use it? When the populations are not normally distributed, the sample size could be small, so, it is very difficult to assemble them as a normal, response is ordinal, that means, integer numbers type response, ok?. So, what do we do? We can say, H_o, that is the null hypothesis is, you are looking at distribution of population A and population B, the null hypothesis could be $A = B$; the Ha, the alternate hypothesis could be one of them, right?; $A \neq B$, that means, population A and population B are not the same; or population A is larger than population B, that is, $A > B$; or population A is less than population B, that is $A < B$. So, these could be the various types of alternate hypotheses and this could be the null hypothesis; just like the previous parametric based test, actually.

(Refer Slide Time: 05:27)



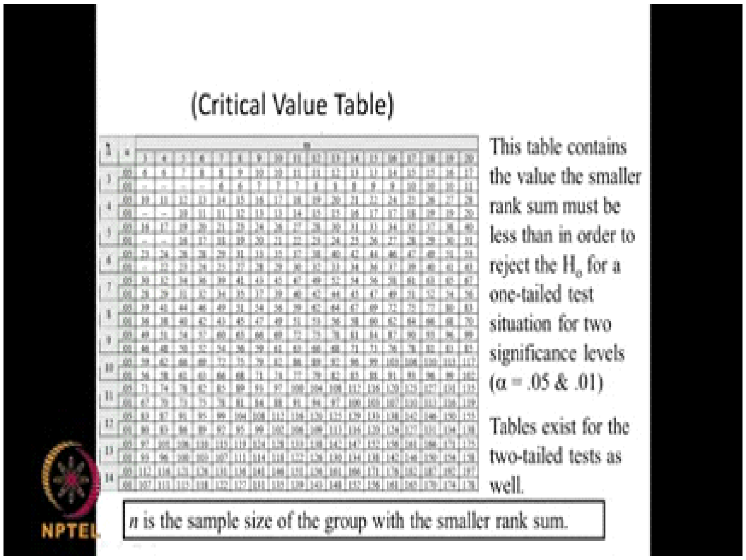
Let us look at some problems later on, but initially, you are looking at A = B; distribution of A and B are same; whereas A ≠ B right, here, A > B, because A is shifted to the right. So, for example, we can say, is there an evidence that the values of A are generally larger than that of B? If this is on this side, we can say, is there evidence that the values of A is generally less than that of population B. These are some good references, which I am talking about here, you know; it is worth looking at these references, books, which gives you quite a lot of examples related to health sciences and biological sciences.

(Refer Slide Time: 06:14)



So, what do we do, in this, we rank all the data. So, suppose, we have nA observations for a sample A, nB observations for sample B; so, we combine all the results together and rank them; that means, 1, 2, 3, 4, smallest value called 1, and so on. And then, put them in ascending order, that means, smallest value at the bottom, that is 1, then the next larger is 2, like that; but, you have to combine both sets. Then, sum the ranks of the observations from population A separately, that will be called W_A; sum the ranks of population B, that will be called W_B; assign average rank to these tied observations. Suppose, if I have a data 16 in population A, and I have another data 16 in population B, what do I do? I will add 2 ranks, ÷ 2; that is called the average rank. So, the alternate hypothesis is, A > B, reject H_o; if W_A is smaller than W_B. So, it is exactly like that. So, if W_A rank is smaller than W_B, then, A < B; that is the alternate hypothesis. And, if W_A is big than W_B, then, alternate will be A > B. And of course, the null hypothesis will be A = B. So, you understand!. So, what we do is, we combine all the data and then rank them; the smallest will get rank 1, then, next one will be rank 2, like that, the largest will get the highest rank. And then, you sum all the ranks related to A, and call it W_A; sum all the ranks related to B, call it W_B. And then, if W_A is smaller than W_B, the alternate will be A < B; if W_A is bigger than W_B, then the alternate will be A > B. So, you understand this?. And then, there is a table for this. There is a table also for ranks, and then, we can reject the null hypothesis, or accept the null hypothesis, based on the statistics, and the table comparison, just like the normal parametric test.

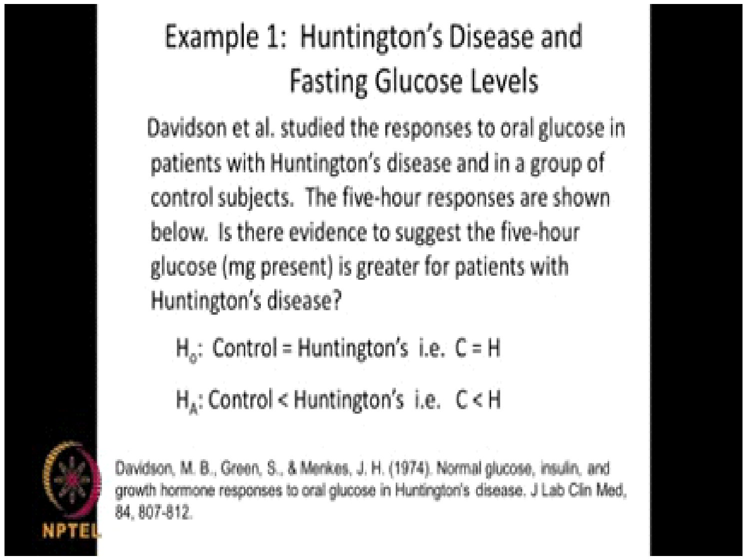
(Refer Slide Time: 08:27)



Let us look at a problem on… Before going to that, this is the table. This is the table, as you can see, n is the sample size of the group with the smaller rank; n is the sample size of the group with the smaller rank; and this here, you have the 2 αs, the 0.05 probability and 0.01 probability. So, n, this is the smaller ranks; sum must be less than, in order to reject the H_o, for a one-tailed test.

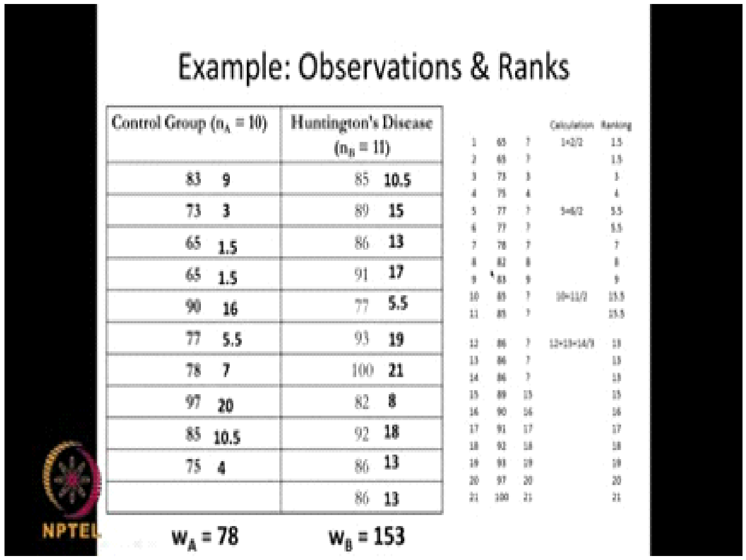
Similarly, you also have a table for the two- tailed test. So, this table contains values; the smaller ranked sum must be less than, in order to reject the H_o for a one-tailed test; do you understand? So, we have here, sample size of the group with the smaller rank sum; and, this is the sample size of the group with the larger rank sum; do you understand?

(Refer Slide Time: 09:24)



Now, let us do a problem here. This problem is taken from this particular references. It is the Journal of Laboratory Clinical Medicine. We are looking at the fasting glucose levels of patients who are having Huntington's Disease, and those of the control. So, glucose is given orally to those with Huntington's disease, as well as to a group of the control; and then, after 5 hours, the data is collected; that is, the glucose present, in terms of mg. The hypothesis is, is the glucose greater for patients with Huntington disease; that means, H_o is control equal to Huntington; we will call it C = H; Ha is, control < Huntington; do you understand? Control < Huntington. So, C < H. So, as I said, this example is nicely taken from this reference.

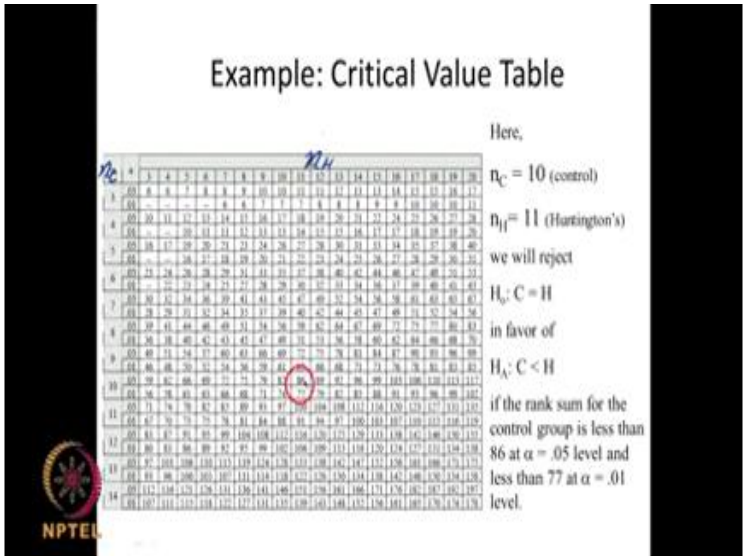
(Refer Slide Time: 10:20)



So, what do you do? This is the data. So, you have the control glucose, 10 control and then, 11 with Huntington; 11 with Huntington, 10 with… This is the glucose levels after 5 hour, after oral injection. This is the value. This is the value. So obviously, we cannot consider this as a normal distribution, because it is sort of ordinal. So, what do we do? We combine all these data; so, 11 + 10, 21 sets of data, and then, we arrange them in the ascending order; that means, increasing order. So, we give 65 starting, then, go right up to 100; this 100 will come there. So, that will get a mark, rank of 21; 65 will… There are two 65s, as you can see, here, here. So, this will get a rank of 1.5, because 1 + 2. So, these two 65s will get 1.4, 1.2 rank. Do you understand? Then comes 73; then comes 75; then comes 77; there are two 77s; here one 77, then here. So, this will get rank of 5 and 6. So, 5 + 6, / two, 5.5, 5.5; like that, you rank all the data. Now, some of them belongs to A, some of them belongs to B, do not forget that. So, we need to separate the As together and the Bs together. And then, add up to get W_A and W_B, ok?.

So, again here, we are putting down these ranks here. So, as you can see here, in this particular case, these two got 1.5 rank. And then, you have the 3 rank and so on, right up to 21. This got the rank 21. So now, if you add up all these ranks, that will get you W control; if you add up all these ranks, we get the W_H. Now, the goal is to tell the null hypothesis is, is the control = H, or control < H; that is the alternate hypothesis. So, we need to go to the table to crosscheck. Let us look at it. W_A is 78, if you add up all these, and W_B is 153, if you add up all these things. So, once you do that, now, we go to the table. This one corresponds to the lower one and that one corresponds to the higher one. So, we look at 10 and 11; so, we get 95 % is 86, and 99 % is 77; so, if the data which you get 78 is smaller, ok?.

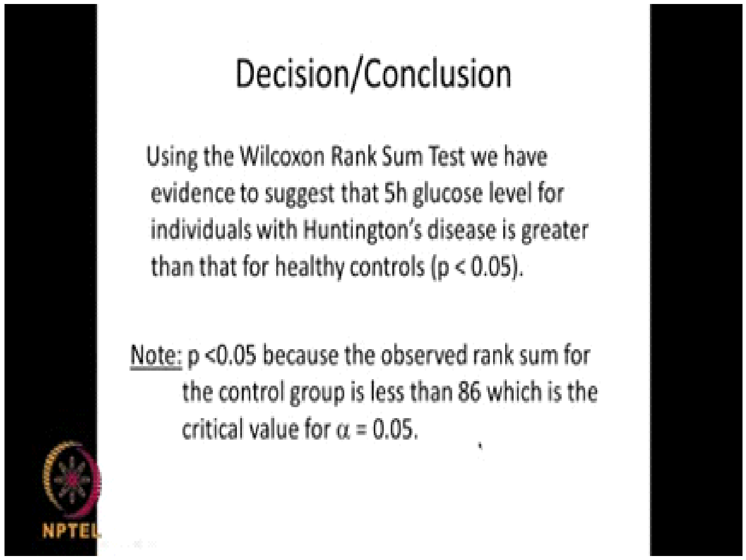
(Refer Slide Time: 12:46)



So, nc is 10; that is, the smallest comes here, and the largest comes here. So, this is 11, this is 10. We will reject if the rank sum of the control group is < this 86, at 95 %, or < 77. So, what do we get the answer is 78; so here; 78 is < 86; so, at 95 %, or α of, or probability of 0.05, we can reject the null hypothesis. So, what is the null hypothesis? The null hypothesis is the blood glucose level is the same for the control as well as those suffering from Huntington disease, and as per this, we got a number of 78, that is the smallest number, and when you look at the smallest number and compare it here, this number is larger, so, we reject the null hypothesis.

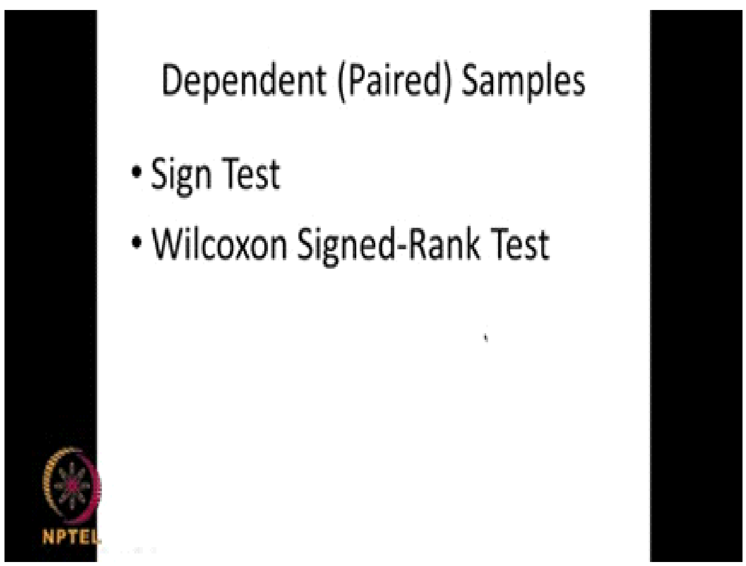
So, we accept this alternate, that is, control will have lower H; do you understand? So, you look at this table, and then, these rows correspond to the lower data points; this corresponds to the higher data points. In this case, this is control and in this case, this is the Huntington disease. So, if you take the 95 %, you get 86; but, when you add up all the control ranks, we get it as 78; because, 78 is lower than 86, we reject the null hypothesis. If that W_A has been larger than 86, then, we can, cannot reject the null hypothesis; but, in this case, we need to reject the null hypothesis. So, that is one problem.

(Refer Slide Time: 14:34)



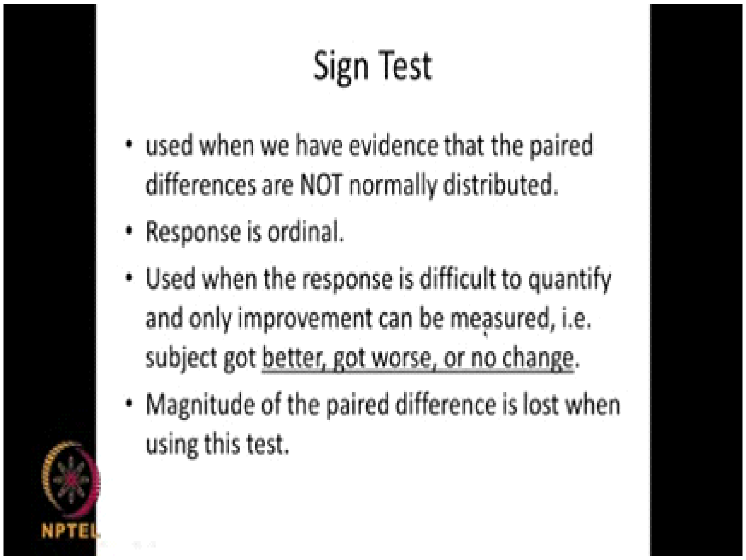
So, using the Wilcoxon Rank Sum Test, we have evidence to suggest that the 5 hour glucose level for individual with Huntington disease is greater than that for the healthy controls at 95 %, or p is ≤ 05. We took this because, the 78 < 86, at the value of this.

(Refer Slide Time: 14:56)



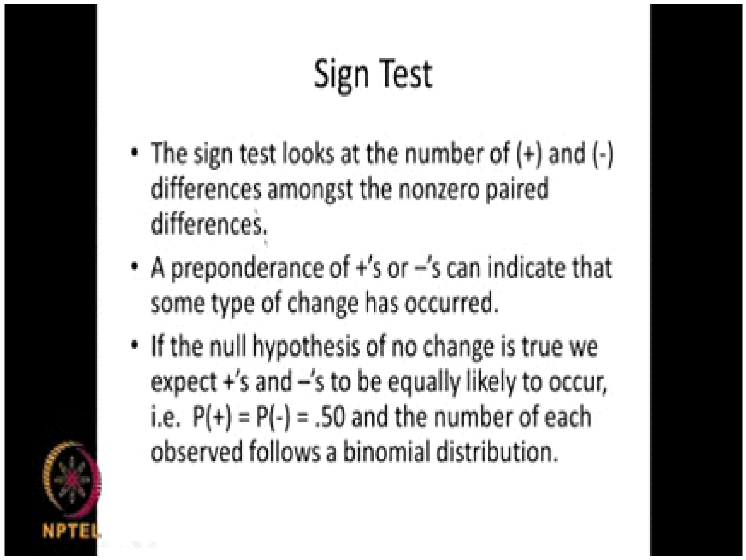
Let us look at the paired sample. So, we looked at two samples, equivalent to two sample t test. We got this. Now, dependent samples, that means, paired t test, where you are comparing, you are using the same subjects for control, as well as drug. There are two different types of test. One is called the Sign test; other is called the Wilcoxon Signed-Rank Test.

(Refer Slide Time: 15:20)



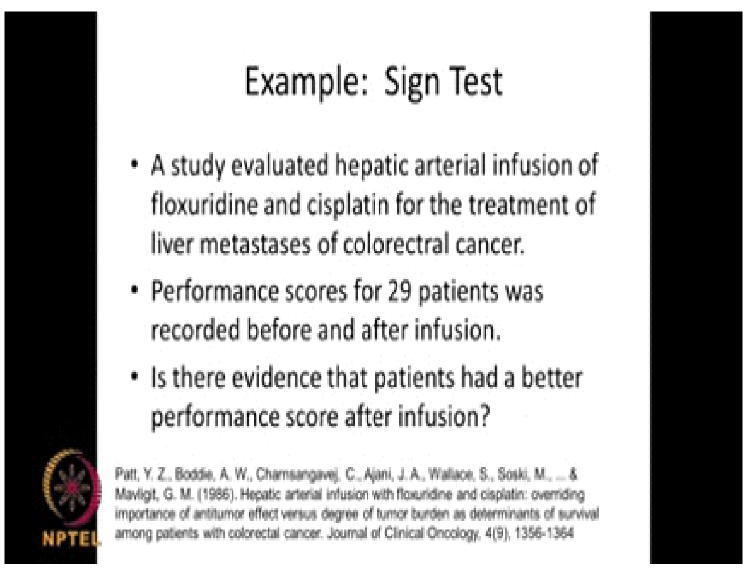
Let us look at the Sign test. So, this is used when we have no evidence that the paired differences are not normally distributed. So, if it is normally distributed, it is not difficult; we can use the paired sample t test, and I taught you how to do that long time back, right?. The response is ordinal; that means, it is just numerics; there is no independent variable, like, as a function of time and so on. Used when the response is difficult to quantify. Sometimes, we may have situations like better, got worse, no change, marginal change, ok?, that sort of thing. But, the main drawback of this test is, the magnitude of the paired difference is lost, because it is looking at only the signs, plus, minus comparing the pluses with the minus; the magnitude, how much it is changed, is lost. But, the other test can help you out of that; the Wilcoxon Signed Rank Test can help you out of that; but the sign test loses it, ok?.

(Refer Slide Time: 16:15)



So, it looks at the pluses, it looks at the minuses, and then, it considers only the nonzero paired differences; if there are 0, you neglect it. So, if you have more of pluses than minus, we can indicate that some type of change has occurred. So, if the null hypothesis is, no change is true, we expect pluses and minus to be equal. So pluses, minus 0.5; it should follow a binomial distribution. But, if it does not follow with the probability of 0.05, then, we can use the test to find out whether it is significant or not, ok?. So, it is quite simple.

(Refer Slide Time: 16:54)



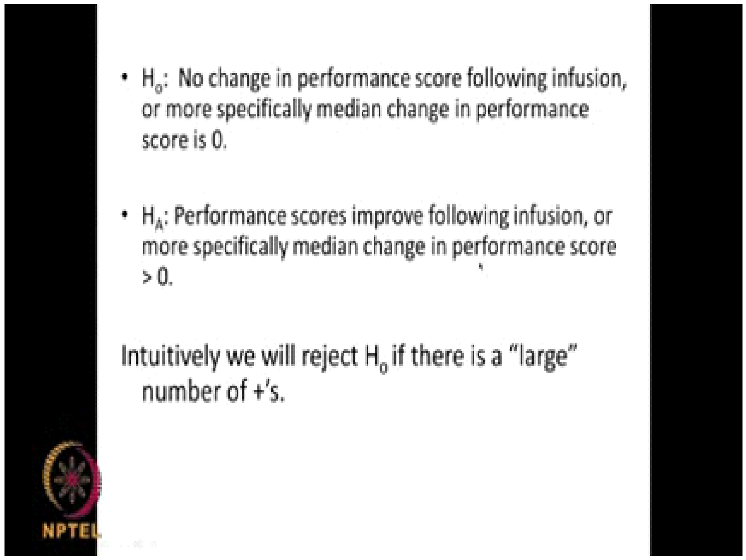
Let us look at an example which uses the sign test. This was taken from this particular book, sorry, reference, Journal of Clinical Oncology. So, they are looking at hepatic arterial infusion of floxuridine and cisplatin. These are anti cancer drug for the treatment of a colorectal cancer. So, 29 patients' data was recorded before and after infusion. So, is there any evidence? So, there was an infusion of floxuridine and cisplatin, anti cancer drugs for the treatment of colorectal cancer. It was given to 29 patients; data was recorded before and after infusion. So, is there evidence that patients had a better performance score after infusion, that is the question here asked.

(Refer Slide Time: 17:47)

Patient	Before (B) Infusion	After (A) Infusion	Difference (A – B)	Patient	Before (B) Infusion	After (A) Infusion	Difference (A – B)
1	2	1	-1	16	0	0	0
2	0	0	0	17	0	3	3
3	0	0	0	18	2	3	1
4	1	0	-1	19	2	3	1
5	3	3	0	20	3	2	-1
6	1	0	-1	21	0	4	4
7	1	3	2	22	0	3	3
8	0	0	0	23	1	2	1
9	0	0	0	24	0	3	3
10	0	0	0	25	0	2	2
11	1	0	-1	26	1	1	0
12	1	1	0	27	3	3	0
13	2	1	-1	28	1	2	1
14	3	1	-2	29	0	2	2
15	0	0	0				

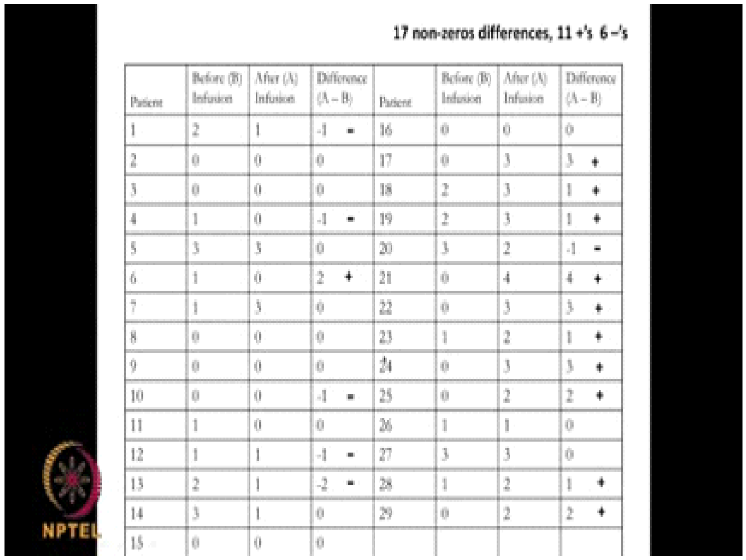
Let us look at the data. So, before infusion, we have there are 29 data points. Look at before infusion, you got this data. This is sort of a qualitative 0, 1, 2, 3; after infusion, same patients, that is why, it is almost like a paired t test which we used to do for parametric; so, you get like this. Now, you subtract (A – B). So, we get negatives, some of them are 0s; so, we get negatives; some of them are 0s; you get some positives as well, and so on, actually, after infusion.

(Refer Slide Time: 18:22)



What do you do? H_0, no change in performance; of course, status quo following infusion, or more specifically, median change in performance score is 0. H_A is performance scores improved following infusion, or more specifically, median change in performance score is > 0, because we are subtracting (A – B). So, we expect it to be better; if it had been (B-A), it would have, that is different. But here, we are doing it after infusion and before infusion; so, subtracting (A – B); so, we expect it to be better; that means, it should be > 0. Let us look at this data. So, there are some 0s; so, forget about that; so, some negatives are there, that means, the after infusion, it is not good. But, there are some positives. So, lot of positives appear to be there. So, we feel that after infusion, things are better, as you can see. Now, let us add up the negatives; let us add up the positives; forget about the zeroes here; do you understand? Intuitively, we may think that infusion has helped, because we have lot of positives coming when compared to the negatives, but we need to do a proper statistical test. Let us look at these data again.

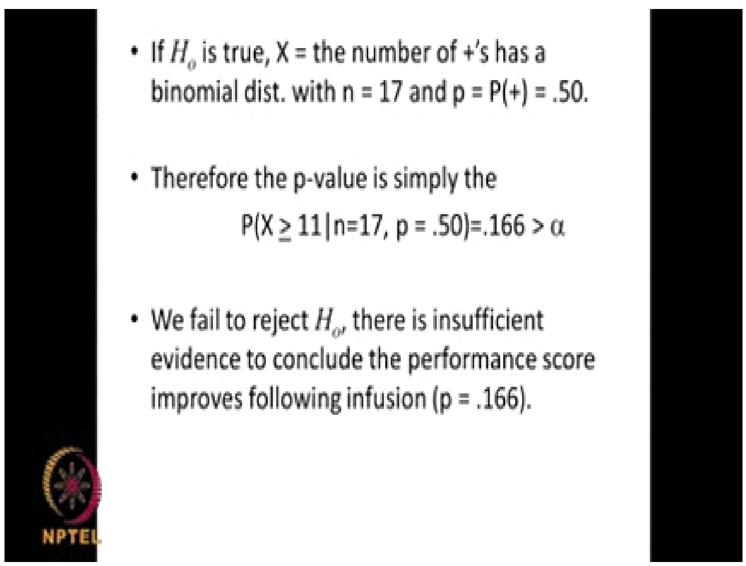
(Refer Slide Time: 19:30)



Patient	Before (B) Infusion	After (A) Infusion	Difference (A – B)	Patient	Before (B) Infusion	After (A) Infusion	Difference (A – B)
1	2	1	-1 -	16	0	0	0
2	0	0	0	17	0	3	3 +
3	0	0	0	18	2	3	1 +
4	1	0	-1 -	19	2	3	1 +
5	3	3	0	20	3	2	-1 -
6	1	0	2 +	21	0	4	4 +
7	1	3	0	22	0	3	3 +
8	0	0	0	23	1	2	1 +
9	0	0	0	24	0	3	3 +
10	0	0	-1 -	25	0	2	2 +
11	1	0	0	26	1	1	0
12	1	1	-1 -	27	3	3	0
13	2	1	-2 -	28	1	2	1 +
14	3	1	0	29	0	2	2 +
15	0	0	0				

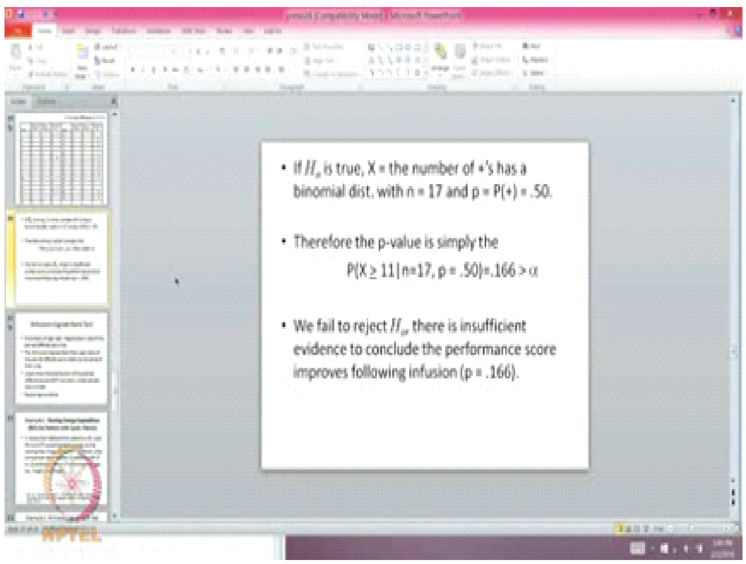
So, we look at only the pluses and minuses. So, as you can see the minus 1, 2, 3, 4, 5, 6; plus is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11; so, you have 11 pluses and 6 minuses in 17 data points.

(Refer Slide Time: 19:57)



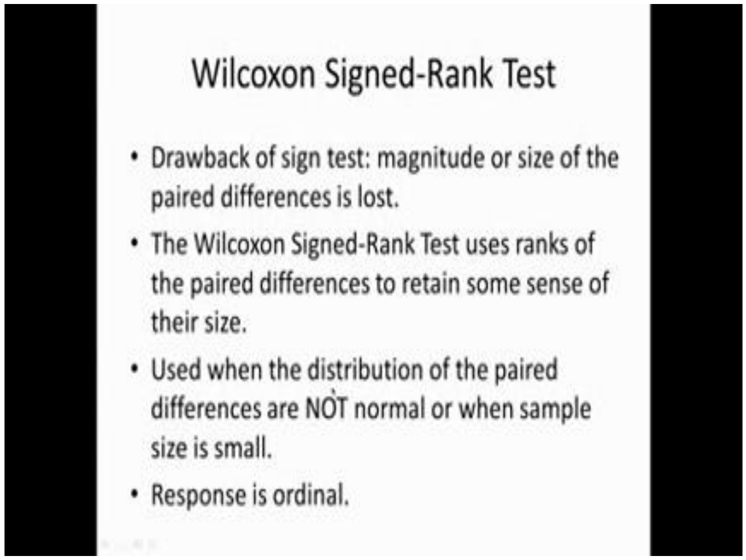
So, if $H_o$is true, then, we should have a probability of 0.5 for the 17 data; whereas, if not, then obviously, we need to use the binomial; out of 17, out of 17 we get 11 positives; out of 17 we get 11 positives; probability is 0.5. So, what will be the p value? So, we failed and this is the p value. But, the p value is much larger, so, we failed to reject the H_o. So, there is insufficient evidence to conclude. Although, it looks lot of positives, but then, 11 out of 17, when you put a p of 0.05, it gives you a probability of 0.166; it gives you a probability of 0.166. Obviously, we cannot reject the null hypothesis. There is insufficient evidence to conclude the performance score. We use a binomial approach to calculate the probability, and let me tell you how to get this particular value. So, the probability of $X \geq 11$, when n is 17, with the probability of 0.5, that is what we need to calculate. So, we need to calculate binomial distribution when 11 successes out of 17, with the probability of 0.5, 12 successes out of 17 with the probability of 0.5, 13 successes out of 17 with the probability of 0.5, 14 successes out of 17 with the probability of 0.5, 15 successes out of 17 with the probability of 0.5, 16 out of 17 with probability of 0.5 and 17 out of 17 with probability of 0.5 and add all of them, and that should give you this.

(Refer Slide Time: 21:52)



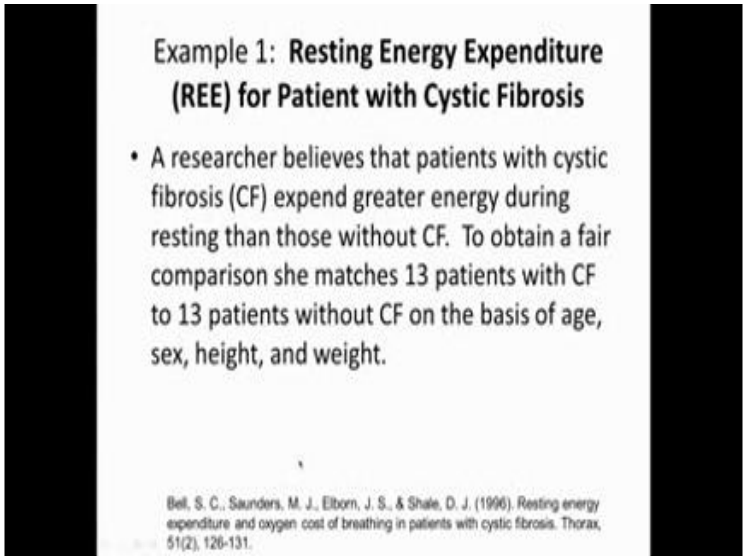
Let us look at it. We use the same binomial distribution; we use the binomial distribution relationship. So, binom, 11 out of 17 with the probability of 0.5, false, that should give you… So, this will become 12 out of 17, and this will become 13 out of 17; this will become 14 out of 17; this will become 15 out of 17; this will become 16 out of 17; and finally, this will become 17 out of 17. So, if you add up all these terms together, you will get 0.166. And, this is how you get this probability. This probability tells you the probability of more than or equal to 11 in a size of 17, if the probability of occurrence is 0.5.

(Refer Slide Time: 23:31)



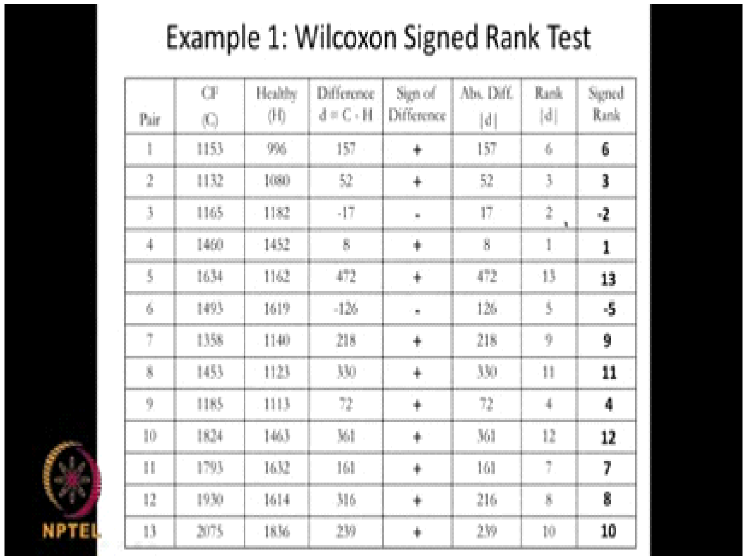
But then, the sign test has this biggest drawback, because it does not consider the magnitude of the sign; it just looks at the signs, plus, minus and so on. So, in order to overcome that particular problem, we have something called Wilcoxon Signed-Rank Test. It does little bit consider the magnitude. Like I said, sign test, the magnitude or the size of the paired differences is lost. So, Wilcoxon Signed Rank Test looks at the sign change, as well as it also considers the magnitude by some sort of a ranking, ok?. Generally, we use it when they are not normal. Obviously, like I said, we are looking at non normal distribution, and the response is ordinal, like the previous case.

(Refer Slide Time: 24:22)



So, let us look at one problem. This is based on a particular reference here. So, resting energy expenditure for patient with Cystic Fibrosis. This is the example taken from this particular reference. So, researcher believes that patient with Cystic Fibrosis expend greater energy during resting than those without CF. So, to obtain a fair comparison, 13 patients with Cystic Fibrosis, 13 patients without Cystic Fibrosis are compared. So, the hypothesis is, CF expends greater energy during resting than those without CF. So, the data was then averaged out, or adjusted for age, sex, height and weight.

(Refer Slide Time: 25:10)

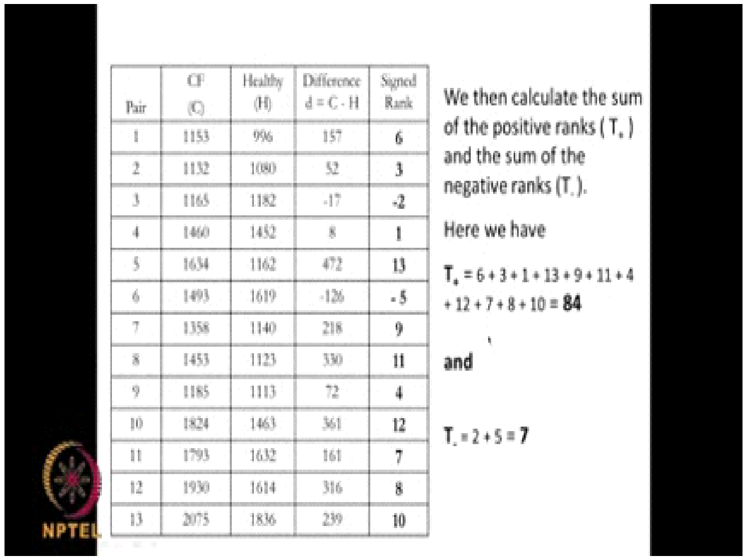


Example 1: Wilcoxon Signed Rank Test

Pair	CF (C)	Healthy (H)	Difference d = C - H	Sign of Difference	Abs. Diff. |d|	Rank |d|	Signed Rank
1	1153	996	157	+	157	6	**6**
2	1132	1080	52	+	52	3	**3**
3	1165	1182	-17	-	17	2	**-2**
4	1460	1452	8	+	8	1	**1**
5	1634	1162	472	+	472	13	**13**
6	1493	1619	-126	-	126	5	**-5**
7	1358	1140	218	+	218	9	**9**
8	1453	1123	330	+	330	11	**11**
9	1185	1113	72	+	72	4	**4**
10	1824	1463	361	+	361	12	**12**
11	1793	1632	161	+	161	7	**7**
12	1930	1614	316	+	216	8	**8**
13	2075	1836	239	+	239	10	**10**

So, the data is given here. Cystic Fibrosis. These are the data, and these are the healthy. So, the comparison is Cystic Fibrosis expends more energy. So, what we have done is C - H, C - H, C - H, H. Again, then look at the sign difference; this is plus, plus, minus, plus, plus, minus, plus, plus, plus, plus, plus, plus and so on, actually. Now, you take the absolute difference of this data, and then, you rank them; just like originally we did rank tests, so the lowest gets 1; the next one gets 2, and so on, until the highest gets 13. What is signed rank? You put the sign also now; so 2 will become - 2; 1 will become + 1 and so on. So, this will become - 5, and rest all will become plus corresponding rank. So, it considers not only some sort of the absolute values, as well as it considers the rank.

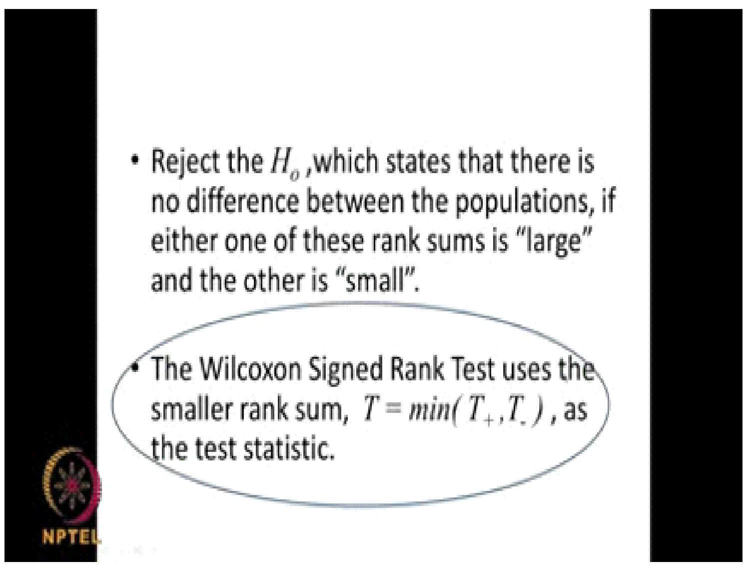
(Refer Slide Time: 26:18)



Pair	CF (C)	Healthy (H)	Difference d = C - H	Signed Rank
1	1153	996	157	6
2	1132	1080	52	3
3	1165	1182	-17	-2
4	1460	1452	8	1
5	1634	1162	472	13
6	1493	1619	-126	- 5
7	1358	1140	218	9
8	1453	1123	330	11
9	1185	1113	72	4
10	1824	1463	361	12
11	1793	1632	161	7
12	1930	1614	316	8
13	2075	1836	239	10

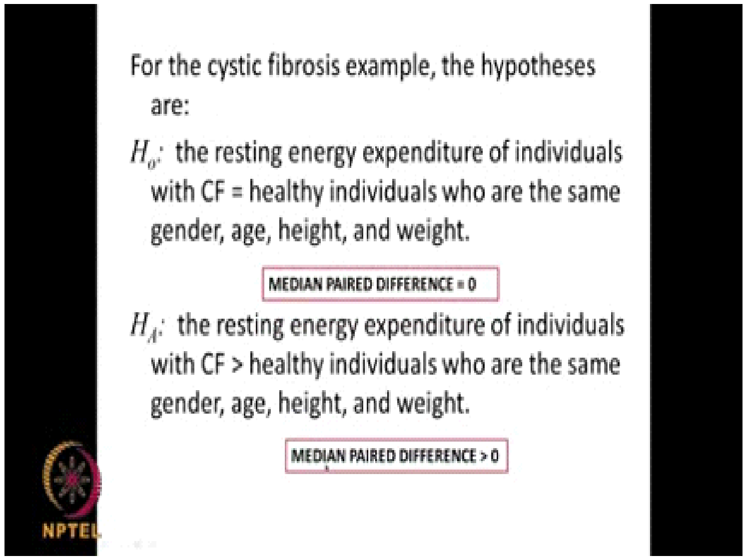
Now, you, so we have this 13 sets of data. We look at them. Then, we calculate the positive ranks, sum of the positive ranks and then, we calculate the sum of the negative ranks. So, negative rank is 2 + 5 is 7; positive ranks, rest of them, T + and T -.

(Refer Slide Time: 26:36)



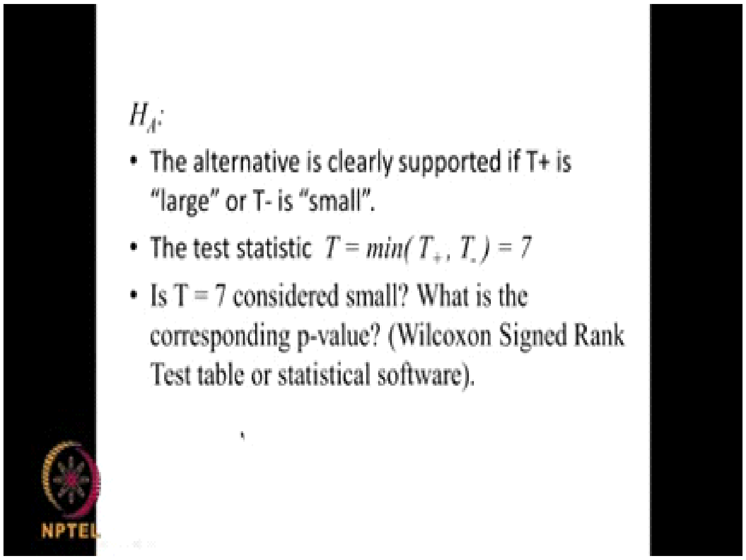
Reject the H_o which states that there is no difference between the populations, if either one of these rank sums is large, and the other is small. So, the Wilcoxon Signed Rank Test, it uses the smaller rank from these 2 ranks, minimum of that row; just like the original rank test, it uses that as the test statistics.

(Refer Slide Time: 26:59)



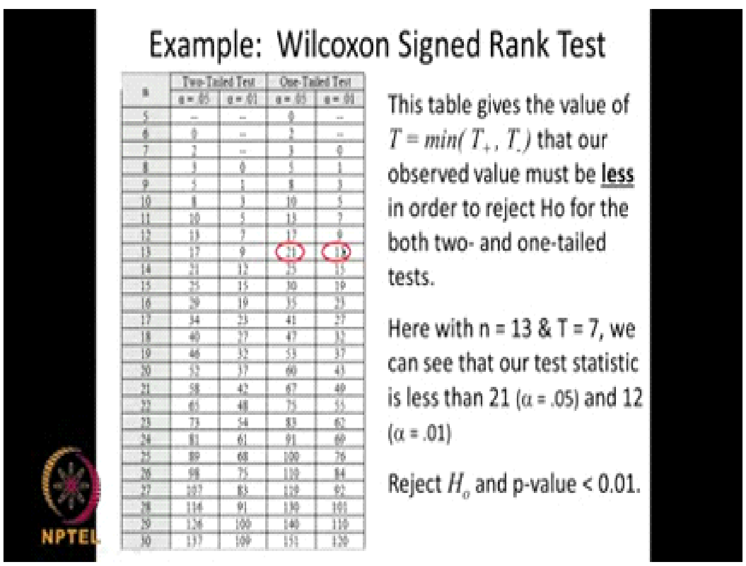
So, the H_o here is the resting energy expenditure of individuals with CF = the healthy individuals who are the same gender, age, height; that means, there is no difference in the energy spent by the CF and non CF. And, H_A, the alternate hypothesis is the resting energy expenditure of individuals with CF > the healthy individuals, after adjusting for gender, age, height and weight. So, we are doing this paired difference; 0 for H_o; > 0 for H_A. Do you understand?

(Refer Slide Time: 27:34)



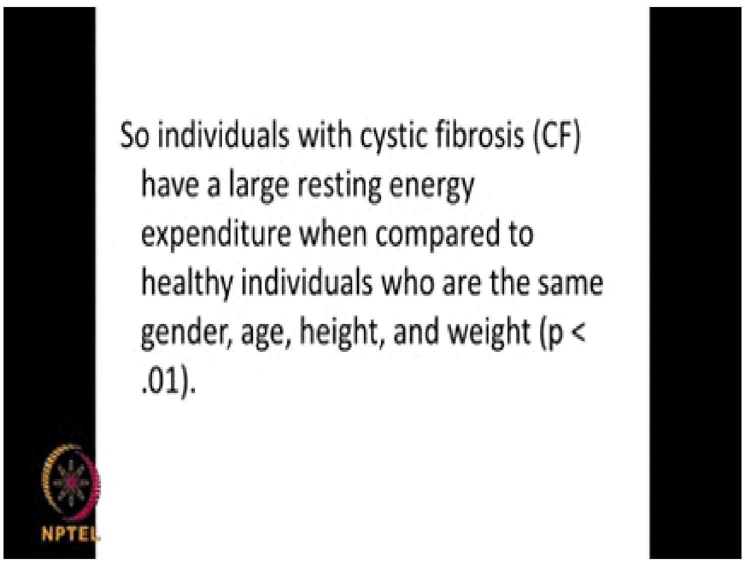
So, we take the smaller of the two, in this case the -, which is 7. Now, is T 7 much smaller, and then calculate what is the p value, using the Wilcoxon Signed Rank Test or statistical.

(Refer Slide Time: 27:50)



So, look at the Wilcoxon Signed Rank Test table. So, we have this degree, the number of data points. We have the two-tailed and one-tailed; let us look at the one-tailed. So, 21 and 12 we get. Observed value is much less than this value, so, we reject … So, out of the 13 and 7… So, we take the 7 here. 7, and the total number of data points is 13. And here, it is 7; it is the data we are talking about. So, we go to the, total number is 13 and we are comparing with 7. So, you should get for 95, or 99, 21 and 12. So, these numbers are much larger than the 7. So, reject H_o at even $p = 0.01$; do you understand? It is exactly like the previous signed rank test we were looking at, the table. Here, the total number, because it is paired, we have only one set of number, whereas, in the other rank test, you remember, we had rows and columns, because we had, it is almost equivalent to your two sample t test. Then, these test statistics must be lower than the table value to reject null hypothesis. So, we got 7 and here, we are having 21, or even 12; so, these are much larger than the test statistics. So, we reject the H_o at a p value of 0.01, do you understand?.

(Refer Slide Time: 29:31)



So, individuals with cystic fibrosis have a larger resting energy expenditure when compared to healthy individuals, at $p < 0.1$. Do you understand? So, how to do this problem. So, we looked at two types of nonparametric test, one for equivalent to two sample t test; the other one is equivalent to a paired sampled t test, and corresponding signed test, rank test. Then, we had the paired rank test and so on, each one of them have their own tables for comparison. So, we will continue on this nonparametric test in the next class also.

Thank you very much.

Key Words: Nonparametric tests, Parametric tests, Null hypothesis, alternate hypothesis, homogeneity of variance, ordinal response, normal response, rank, sign, signed rank, rank sum