Transcript

Video – 0424 Unit4B Case CQ Two Independent Samples B

01. 00:01 / 00:07 - Quickly we will run through the steps in hypothesis testing here for the two-sample t-test. The
02. 00:07 / 00:11 - hypotheses are similar to what we've discussed before except now we're comparing two groups
03. 00:11 / 00:17 - very literally and they're not paired, we cannot reduce them to one sample. So we have
04. 00:17 / 00:21 - to contemplate how to compare these two samples. So our null hypothesis is always going to
05. 00:21 / 00:28 - be that the difference between the population means is zero. This is the same as saying
06. 00:28 / 00:34 - that the population mean from group 1 (mu_1) is equal to the population mean from group
07. 00:34 / 00:41 - 2 which we call mu_2. The alternative is one of the three common choices which is either
08. 00:41 / 00:46 - mu_1 minus mu_2 is different from zero, for two-sided, greater than zero, or less than
09. 00:46 / 00:51 - zero for one sided. For the most part when we carry out entire tests, we are going to
10. 00:51 / 00:58 - stick with the two-sided scenario. That is almost always done in practice. And then we
11. 00:58 / 01:03 - can use our confidence interval to tell us whether we believe that the difference is
12. 01:03 / 01:08 - actually positive or negative. But when we initially approach the test, when we set up
13. 01:08 / 01:13 - our hypotheses to begin with, it is most common that we will just ask, is there a relationship
14. 01:13 / 01:18 - and then after we know there is a relationship, we will go talk a little bit about what we
15. 01:18 / 01:25 - see about that relationship. So again the null hypothesis is that mu-1 minus mu_2 is
16. 01:26 / 01:31 - zero versus the alternative hypothesis, most often will be that mu_1 minus mu_2 is different
17. 01:31 / 01:37 - from zero, when we conduct the entire test. But we could ask you to set up some hypotheses
18. 01:37 / 01:42 - where it is going to be we are interested in greater than or we are interested in less
19. 01:42 / 01:47 - than, in that we want to show a specific kind of effect, we want to show our treatment is
20. 01:47 / 01:53 - better than the previous treatment. But again in practice most often the p-value that you
21. 01:53 / 01:59 - will see reported even for cases where you want to prove one directional hypotheses,
22. 01:59 / 02:06 - you almost always still see the two-tailed hypothesis p-value along with a confidence
23. 02:06 / 02:10 - interval that then tells you whether it was greater or less and whether that supported
24. 02:10 / 02:17 - your original research question. The null hypothesis can also be thought of as saying there's
25. 02:17 / 02:22 - no relationship between our quantitative response variable Y and our categorical explanatory
26. 02:22 / 02:29 - variable X. And the alternative hypothesis, when we say not equal to, is that there is
27. 02:29 / 02:34 - some relationship between these two variables. So we can always, in all the later modules,
28. 02:34 / 02:40 - no matter whether we are in case CQ, CC, or QQ, our hypotheses can almost always be written
29. 02:40 / 02:46 - as: there is no relationship between the two variables under study versus there is a relationship
30. 02:46 / 02:50 - between the two variables under study and then it is just the kind of test that we do
31. 02:50 / 02:56 - in each of those scenarios. So our next step is going to be to obtain our data, check our
32. 02:56 / 03:01 - conditions, and summarize the data. The two-sample t-test can be used as long as the two samples
33. 03:01 / 03:07 - are indeed independent. Be very careful that you are not using two sample methods when
34. 03:07 / 03:13 - in fact you have paired data. Ask yourself, is the observation in this group paired with
35. 03:13 / 03:18 - an observation in the other group? If the answer is yes, this is not the correct test
36. 03:18 / 03:24 - to do. And we also have to be in one of the following scenarios based upon whether our
37. 03:24 / 03:30 - populations are normal or we have a large enough sample. Very similar to what we discussed
38. 03:30 / 03:37 - in our other t-tests earlier. So both populations are normal or more specifically the distribution
39. 03:37 / 03:43 - of the response variable Y in both populations is normal, since sometimes you might think
40. 03:43 / 03:48 - about the populations as people, well people aren't normally distributed, but the response
41. 03:48 / 03:54 - that we're measuring on these people is normally distributed. Both samples are random or at
42. 03:54 / 03:59 - least can be considered as a random. And in practice checking normality is done the same
43. 03:59 / 04:04 - way we've discussed before, looking at histograms, looking at normal probability plots, and then
44. 04:04 / 04:11 - looking for signs of extreme skewness or extreme outliers. If the populations are not normal
45. 04:11 / 04:16 - that's not necessarily the end of the road because if our samples are large enough then
46. 04:16 / 04:21 - we can really use these methods anyway. So if the populations are known or are discovered

47. 04:21 / 04:27 - not to be normal but the sample size of the random samples is large enough, in our case
48. 04:27 / 04:31 - we will say if it is greater than 30, it is large enough. In practice the larger the sample
49. 04:31 / 04:36 - the better you're going to be in terms of the results. So once we have looked at our
50. 04:36 / 04:42 - conditions and we've decided that we can proceed, we still have one additional issue with the
51. 04:42 / 04:48 - two-sample t-test. Before we go into these two choices let's first talk about some notation.
52. 04:48 / 04:54 - So we are going to have two populations. Each of those populations we are going to have
53. 04:54 / 04:59 - a sample from that population. So we have two samples. Each of those has a sample size.
54. 04:59 / 05:05 - So we will call those n_1 and 2. Each of those samples also has a sample mean, so we will
55. 05:05 / 05:11 - call that Y_1-bar and Y_2-bar. And each has a sample standard deviation which we will
56. 05:11 / 05:17 - call s_1 and s_2. And then we'll have one additional notation that we will introduce
57. 05:17 / 05:22 - in a little while, which is the pooled estimate of the population standard deviation which
58. 05:22 / 05:28 - we create whenever we assume that there are equal variances in each of the two populations
59. 05:28 / 05:35 - which is one of our choices. So here are the two cases for our test statistic. In the first
60. 05:36 / 05:42 - case we have the assumption of equal variances. If it is safe to assume that the two populations
61. 05:42 / 05:49 - have equal variances or equal standard deviations then we can pool our estimates of s_1 and
62. 05:50 / 05:55 - s_2 into some common estimate of the standard deviation. And then we have the following
63. 05:55 / 06:01 - test statistic. We are not going to discuss the details here except that you have access
64. 06:01 / 06:06 - to what the formula is and a little bit of discussion in terms of where it is coming
65. 06:06 / 06:13 - from. So the idea is we're estimating a common population standard deviation, both groups
66. 06:13 / 06:18 - come from populations with the same standard deviation. We have two samples. We want to
67. 06:18 / 06:23 - use those to estimate what we call a pooled estimate of the standard deviation, which
68. 06:23 / 06:30 - is this S_p formula at the bottom. The idea of the test statistic is very similar to what
69. 06:30 / 06:37 - we've done before. On top we have an estimator of our quantity, which is mu_1 minus mu_2.
70. 06:37 / 06:44 - Our estimator is Y_1-bar minus Y_2-bar minus our null value divided by a standard error.
71. 06:44 / 06:48 - So that's what we really want to know about this test statistic, is that it's always going to be our
72. 06:48 / 06:55 - estimator minus our null value divided by our standard error unless we tell you otherwise.
73. 06:56 / 07:01 - There are some tests that aren't formulated this way but at the moment this is what we're
74. 07:01 / 07:05 - looking at in all of the tests that we've seen so far. The second case is the case of
75. 07:05 / 07:10 - unequal variances and this we will use this anytime we have found that it's not safe to
76. 07:10 / 07:16 - assume the two populations have equal standard deviations or equal variances. So we have
77. 07:16 / 07:21 - unequal standard deviations in this case and we use a slightly different t-statistic, but
78. 07:21 / 07:27 - it is still the same idea. On top we have an estimator, Y_1-bar minus Y_2-bar minus
79. 07:27 / 07:32 - a null value divided by the standard error. But the standard error is slightly different
80. 07:32 / 07:37 - depending on whether we have equal variances or unequal variances and the reasons for that
81. 07:37 / 07:43 - are more statistically related, more mathematically related, and they're certainly not something
82. 07:43 / 07:48 - that we really need to understand beyond the basic idea. We've already developed that the
83. 07:48 / 07:53 - idea behind hypothesis testing is we find an estimator, we have a null value, and we
84. 07:53 / 08:00 - have a standard error. This gives us how many standard errors away from our null value our
85. 08:00 / 08:06 - data were and then we use that to find a p-value which helps us determine whether what we've
86. 08:06 / 08:12 - seen is rare enough to reject the null hypothesis or not. We won't be calculating any of these
87. 08:12 / 08:17 - test statistics by hand but we will rely on software to obtain the value for us. You should
88. 08:17 / 08:22 - be able to locate the value of the test statistic and the p-value, and the confidence intervals
89. 08:22 / 08:29 - that we discuss, in the output from the software. Again remember all the tests so far that we've
90. 08:29 / 08:35 - discussed all have this form they are the test statistic as the estimator minus the
91. 08:35 / 08:41 - null value all divided by the standard error of the estimator that we're using. In this
92. 08:41 / 08:47 - case the p-value relies on a certain t-distribution. In the case where we can assume equal variances
93. 08:47 / 08:52 - the degrees of freedom are actually fairly simple to calculate, they are n_1 plus n_2
94. 08:52 / 08:58 - minus 2. For the case of unequal variances, the formula for degrees of freedom is more
95. 08:58 / 09:03 - complex and we will rely on software really to obtain the degrees of freedom in both cases
96. 09:03 / 09:07 - and provide us with the correct p-value. But there's always going to be this distribution
97. 09:07 / 09:12 - in the background that we are using to look up our p-values when we conduct hypothesis

98. 09:12 / 09:18 - test. The conclusions will be the same as we always have. If the p-value is small then
99. 09:18 / 09:23 - there is a significant difference. In our case this is going to translate into there
100. 09:23 / 09:30 - is a relationship between our quantitative response variable Y and our categorical explanatory
101. 09:30 / 09:36 - variable X. If the p-value is not small then we don't have enough evidence to reject the
102. 09:36 / 09:42 - null hypothesis. We can also use confidence intervals in this case similar to how we did
103. 09:42 / 09:48 - this in the paired t-test example. If the null value of zero falls outside of the confidence
104. 09:48 / 09:53 - interval, we can reject the null hypothesis. If the null value of zero falls inside the
105. 09:53 / 09:59 - confidence interval, then the null hypothesis is not rejected. So that's the basic process
106. 09:59 / 10:03 - of the t-test. You will see as we go through the remaining material we're going to try
107. 10:03 / 10:08 - to get through the background of these tests very quickly and get into the software to
108. 10:08 / 10:14 - see how do we conduct these tests and what do they mean. So in this case we have to decide
109. 10:14 / 10:19 - whether we believe the variances are equal or not. There is a test for this. You can
110. 10:19 / 10:25 - also just use your common sense and look at the distributions and see if it's reasonable
111. 10:25 / 10:30 - that the variances in the two populations are the same. If it is really ridiculously
112. 10:30 / 10:34 - different, you have a huge variation in one group and a tiny variation in another group,
113. 10:34 / 10:40 - you'll be able to make this decision without any assistance from any hypothesis test. But
114. 10:40 / 10:45 - there is a test called Levine's test for equality of variances and most software packages give
115. 10:45 / 10:52 - the output of this test when you ask for a two-sample t-test. The null hypothesis of
116. 10:52 / 10:58 - this test, the test for equality of variances, is that the standard deviation in group 1
117. 10:58 / 11:05 - is equal to the standard deviation in group 2 versus the alternative hypothesis is that
118. 11:05 / 11:08 - they are different. So you can see that if we have a small p-value then we're going to
119. 11:08 / 11:15 - reject the null hypothesis and conclude that indeed the variances in the two groups, or
120. 11:15 / 11:19 - the standard deviations in the two populations, are different. In that case were going to
121. 11:19 / 11:25 - use case B and have unequal variances. If we fail to reject the null hypothesis here,
122. 11:25 / 11:31 - we have a large p-value, and then we're going to stick with the assumption of equal variances
123. 11:31 / 11:36 - unless we have clear evidence from the distributions that we don't believe that. So you can
124. 11:36 / 11:42 - always override this test by your common sense but this is a method and you will generally use
125. 11:42 / 11:48 - this test to decide which of the two t-tests you're going to use. So again, if you reject
126. 11:48 / 11:54 - the null hypothesis for this test you are saying the variances are unequal. If you fail
127. 11:54 / 12:00 - to reject the null hypothesis you are basically unable to say the variances differ so you're
128. 12:00 / 12:04 - going to be left with the best choice which is to go ahead and assume that the variances
129. 12:04 / 12:10 - or equal or reasonably equal. We will look at the output for this when we get into our
130. 12:10 / 12:11 - example in the software.