1
00:00:00,000 --> 00:00:00,866

2
00:00:00,866 --> 00:00:05,866
So when we reject the null hypothesis in an
ANOVA F-test, basically all we can conclude

3
00:00:05,866 --> 00:00:11,932
is that not all of the means are equal or
we could say there are some differences
between

4
00:00:11,933 --> 00:00:18,499
the means or we could say that the response
Y is related to the explanatory X.

5
00:00:18,500 --> 00:00:25,200
It does not provide any immediate insight
into why we've rejected the null hypothesis.

6
00:00:25,200 --> 00:00:36,900
As a quick exploratory, you could look at
the confidence intervals for each of the groups

7
00:00:36,900 --> 00:00:36,933
individually and we could do this just by
using methods that we've already discussed.

8
00:00:36,933 --> 00:00:41,999
So just finding a one-sample confidence interval
for each of the groups and then comparing

9
00:00:42,000 --> 00:00:43,400
them.

10
00:00:43,400 --> 00:00:48,400
When we compare multiple groups using
standard
95 percent confidence intervals, we have an

11
00:00:48,400 --> 00:00:54,766

increased chance of making an overall error
as each interval has a 5 percent error
individually.

12
00:00:54,766 --> 00:01:00,199
The more comparisons we make the larger
resulting
overall error.

13
00:01:00,200 --> 00:01:05,200
So the idea is if we claim two means are
different,
we may be making a mistake more often than

14
00:01:05,200 --> 00:01:10,700
5 percent of the time if we don't make an
adjustment for the fact that we're doing multiple

15
00:01:10,700 --> 00:01:11,700
comparisons.

16
00:01:11,700 --> 00:01:16,900
So in analysis of variance, one common method
is to use a multiple comparison procedure,

17
00:01:16,900 --> 00:01:19,866
and this will be fairly easy to do in the
software.

18
00:01:19,866 --> 00:01:25,299
The idea is to adjust this error rate so that
overall, we've made a five percent error in

19
00:01:25,300 --> 00:01:28,466
all of our comparisons simultaneously.

20
00:01:28,466 --> 00:01:33,399
So we will look at a few of these in the software
when we look at our example.

21
00:01:33,433 --> 00:01:39,399
So our example we're going to talk about first

is academic frustration related to major.

22
00:01:39,400 --> 00:01:44,133
A college Dean believes that students with
different majors may experience different

23
00:01:44,133 --> 00:01:46,766
levels of academic frustration.

24
00:01:46,766 --> 00:01:52,832
Random samples of 35 business, English,
math,
and psychology majors are asked to rate their

25
00:01:52,833 --> 00:01:59,399
level of academic frustration on a scale from
1 which is the lowest level to 20 which is

26
00:01:59,400 --> 00:02:01,066
the highest level.

27
00:02:01,066 --> 00:02:04,132
We have an image here that illustrates this
problem.

28
00:02:04,133 --> 00:02:05,633
Our major which is X.

29
00:02:05,633 --> 00:02:11,766
We have four majors: business majors, English
majors, math majors, and psychology majors.

30
00:02:11,766 --> 00:02:16,699
We have these four populations representing
all business majors, English majors, math

31
00:02:16,700 --> 00:02:22,666
majors, or psychology majors at this university
and we're going to try to compare the means

32
00:02:22,666 --> 00:02:26,532

between these four groups.

33
00:02:26,533 --> 00:02:28,699
We begin by stating our hypotheses.

34
00:02:28,700 --> 00:02:33,633
The only thing we can do to make this a little
more specific than what we had already is

35
00:02:33,633 --> 00:02:35,733
that we can say how many means there are.

36
00:02:35,733 --> 00:02:38,699
So the null hypothesis is that mu_1 is equal

37
00:02:38,700 --> 00:02:45,766
to mu_2 is equal to mu_3 is equal t o mu_4.
You could also use some notation that made

38
00:02:45,766 --> 00:02:50,432
sense to you, for example, we might use B
for business and M for math and P for
psychology

39
00:02:50,433 --> 00:02:56,133
and so on. The null hypothesis could also
be stated that there is no relationship between

40
00:02:56,133 --> 00:03:01,433
major and academic frustration level. The
alternative hypothesis is that not all mu's

41
00:03:01,433 --> 00:03:07,433
are equal. We could also say that there is
a relationship between major and academic

42
00:03:07,433 --> 00:03:13,466
frustration level. So next we're going to
obtain the data, check our conditions and

43
00:03:13,466 --> 00:03:18,932

summarize the data. In our data we do have all our conditions satisfied. The samples

44
00:03:18,933 --> 00:03:23,799
were all randomly chosen and therefore they are independent. The sample sizes are large

45
00:03:23,800 --> 00:03:29,866
enough in that we have 35 in each group. So we don't really have to worry too much about

46
00:03:29,866 --> 00:03:35,899
normality. We have our boxplots here so we can look at our data more carefully. The data

47
00:03:35,900 --> 00:03:40,666
do suggest that the frustration levels of the business students is generally lower than

48
00:03:40,666 --> 00:03:45,532
the other three majors and then we might want to see if there's any other differences. It's

49
00:03:45,533 --> 00:03:49,799
pretty clear from looking at this graph that I'm probably not going to be able to find

50
00:03:49,800 --> 00:03:55,166
a significant difference between mathematics and psychology but I might be able to find

51
00:03:55,166 --> 00:04:01,166
a significant difference between English and either mathematics or psychology or both depending

52
00:04:01,166 --> 00:04:07,599
on the final results. To determine if we can have equal variances right now, we're just

53
00:04:07,600 --> 00:04:12,466
going to look at the rule of thumb. The rule

of thumb is satisfied since we have the largest

54
00:04:12,466 --> 00:04:21,866
standard deviation is 3.082 and the smallest
is 2.088 and if we divide that we will get

55
00:04:21,866 --> 00:04:29,399
less than 2 and so we can continue with the
test. Here we are using the minitab output

56
00:04:29,400 --> 00:04:34,366
to display the results of the ANOVA. It will
be fairly similar in most software packages,

57
00:04:34,366 --> 00:04:40,132
especially the top table. We have our sources
of variation due to major. That's the variation

58
00:04:40,133 --> 00:04:44,966
between our sample means, which in this case
represents a difference in the means due to

59
00:04:44,966 --> 00:04:52,732
major and then we have our error, which is
the overall variation within each group and

60
00:04:52,733 --> 00:05:01,499
we see an F-statistic of 46.6 and a p-value
of basically 0.000. There is some other
information

61
00:05:01,500 --> 00:05:07,000
contained here but that's all we're going
to talk about right now. So here we have our

62
00:05:07,000 --> 00:05:11,366
- value is very small, basically nothing,
which tells us that it would be next to
impossible

63
00:05:11,366 --> 00:05:16,532
to get data like those observed had the mean

frustration level of the four majors been

64
00:05:16,533 --> 00:05:23,399
the same in the populations, as the null
hypothesis
claims. So in our example, the p-value is

65
00:05:23,400 --> 00:05:28,666
small. That indicates that our data provides
extremely strong evidence to reject the null

66
00:05:28,666 --> 00:05:34,299
hypothesis. We can conclude that the
population
mean frustration level of the four majors

67
00:05:34,300 --> 00:05:40,766
are not all the same. Or in other words, that
majors do have an effect on students' academic

68
00:05:40,766 --> 00:05:47,399
frustration levels at the school where this
test was conducted. And as we mentioned, we

69
00:05:47,400 --> 00:05:52,166
can construct confidence intervals through
standard methods. Here in the minitab output,

70
00:05:52,166 --> 00:05:56,532
we have individual 95 percent confidence
intervals.
These would be what we would obtain if we

71
00:05:56,533 --> 00:06:02,666
just asked for a standard confidence interval
within each group and it is pretty clear without

72
00:06:02,666 --> 00:06:08,099
even having any adjustment for multiple
comparisons
that the business students have a much lower

73

00:06:08,100 --> 00:06:13,933
frustration level than the other three groups.
We can see that the group for English does

74
00:06:13,933 --> 00:06:19,266
not overlap with the group for psychology
in terms of the 95 percent confidence interval

75
00:06:19,266 --> 00:06:25,166
for the mean, but again one of the problems
is we have a higher chance of making an error

76
00:06:25,166 --> 00:06:30,732
in one of our comparisons due to the fact
that we have a five percent error rate in

77
00:06:30,733 --> 00:06:35,433
each of our confidence intervals. We will
look at the multiple comparisons once we go

78
00:06:35,433 --> 00:06:36,733
into the software.