Statistical Significance
Section 15.6
Lecture 29

Robb T. Koether

Hampden-Sydney College

Fri, Mar 4, 2016

Outline

Outline

Application

Example (Dr. Koether's Smart Pill)

- I have developed a pill which, if taken the night before a test, will improve a student's performance on the test the next day by at least 5 points.

Application

Example (Dr. Koether's Smart Pill)

- I have developed a pill which, if taken the night before a test, will improve a student's performance on the test the next day by at least 5 points.
- I will test my pill on 6 subjects: A, B, C, D, E, and F.

Application

Example (Dr. Koether's Smart Pill)

- I have developed a pill which, if taken the night before a test, will improve a student's performance on the test the next day by at least 5 points.
- I will test my pill on 6 subjects: A, B, C, D, E, and F.
- Three of them will be assigned at random to the treatment group and will receive the pill.

Application

Example (Dr. Koether's Smart Pill)

- I have developed a pill which, if taken the night before a test, will improve a student's performance on the test the next day by at least 5 points.
- I will test my pill on 6 subjects: A, B, C, D, E, and F.
- Three of them will be assigned at random to the treatment group and will receive the pill.
- The other three will be in the control group and will receive a placebo.

Application

Example (Dr. Koether's Smart Pill)

- I have developed a pill which, if taken the night before a test, will improve a student's performance on the test the next day by at least 5 points.
- I will test my pill on 6 subjects: A, B, C, D, E, and F.
- Three of them will be assigned at random to the treatment group and will receive the pill.
- The other three will be in the control group and will receive a placebo.
- All 6 subjects will take the test the next day.

Application

Example (Dr. Koether's Smart Pill)

- We measure their improvement in test score over the previous test.

Application

Example (Dr. Koether's Smart Pill)

- We measure their improvement in test score over the previous test.
- This test could have been easier or harder than the previous test, which would affect the results.

Application

Example (Dr. Koether's Smart Pill)

- We measure their improvement in test score over the previous test.
- This test could have been easier or harder than the previous test, which would affect the results.
- The control group's scores will be used to determine the increase (or decrease) in score that we would expect based on all factors other than Dr. Koether's Smart Pill.

Application

Example (Dr. Koether's Smart Pill)

- We measure their improvement in test score over the previous test.
- This test could have been easier or harder than the previous test, which would affect the results.
- The control group's scores will be used to determine the increase (or decrease) in score that we would expect based on all factors other than Dr. Koether's Smart Pill.
- The purpose of the experiment is to show that the treatment group had a *greater* increase in test score than did the control group.

Application

Example (Dr. Koether's Smart Pill)

- Here are the increases in test score for the 6 subjects.

A	B	C	D	E	F
5	2	9	-2	0	4

- Whether these results prove the effectiveness of the pill depends on who was in each group.

Application

Example (Dr. Koether's Smart Pill)

- There are 20 possible choices for 3 of the 6 to be in the treatment group.

Treatment	Treatment	Treatment	Treatment
A, B, C	A, C, E	B, C, D	B, E, F
A, B, D	A, C, F	B, C, E	C, D, E
A, B, E	A, D, E	B, C, F	C, D, F
A, B, F	A, D, F	B, D, E	C, E, F
A, C, D	A, E, F	B, D, F	D, E, F

Application

Example (Dr. Koether's Smart Pill)

- Here are the results for the first 10 possibilities.

Treat	Avg	Control	Avg	Increase
A, B, C	5.3	D, E, F	0.7	4.6
A, B, D	1.7	C, E, F	4.3	-2.6
A, B, E	2.3	C, D, F	3.7	-1.4
A, B, F	3.7	C, D, E	2.3	1.4
A, C, D	4.0	B, E, F	2.0	2.0
A, C, E	4.7	B, D, F	1.3	3.4
A, C, F	6.0	B, D, E	0.0	6.0
A, D, E	1.0	B, C, F	5.0	-4.0
A, D, F	2.3	B, C, E	3.7	-1.4
A, E, F	3.0	B, C, D	3.0	0.0

Application

Example (Dr. Koether's Smart Pill)

- Here are the results for the other 10 possibilities.

Treat	Avg	Control	Avg	Increase
D, E, F	0.7	A, B, C	5.3	-4.6
C, E, F	4.3	A, B, D	1.7	2.6
C, D, F	3.7	A, B, E	2.3	1.4
C, D, E	2.3	A, B, F	3.7	-1.4
B, E, F	2.0	A, C, D	4.0	-2.0
B, D, F	1.3	A, C, E	4.7	-3.4
B, D, E	0.0	A, C, F	6.0	-6.0
B, C, F	5.0	A, D, E	1.0	4.0
B, C, E	3.7	A, D, F	2.3	1.4
B, C, D	3.0	A, E, F	3.0	0.0

Application

Example (Dr. Koether's Smart Pill)

- If Dr. Koether's Smart Pill actually has no effect, what is the probability that this experiment will show that it raises test scores by at least 5 points?

Application

Example (Dr. Koether's Smart Pill)

- If Dr. Koether's Smart Pill actually has no effect, what is the probability that this experiment will show that it raises test scores by at least 5 points?
- The probability is 1 out of 20, or 5% chance.

Application

Example (Dr. Koether's Smart Pill)

- What if there had been 12 subjects with 6 people in each group and with the same 6 scores as before (each score occurring twice)?

Application

Example (Dr. Koether's Smart Pill)

- What if there had been 12 subjects with 6 people in each group and with the same 6 scores as before (each score occurring twice)?
- In this case, there are 924 ways to assign 6 of the 12 subjects to the treatment group.

Application

Example (Dr. Koether's Smart Pill)

- What if there had been 12 subjects with 6 people in each group and with the same 6 scores as before (each score occurring twice)?
- In this case, there are 924 ways to assign 6 of the 12 subjects to the treatment group.
- Of the 924 ways, only 31 of them will result in a difference of at least 5 points.

Application

Example (Dr. Koether's Smart Pill)

- What if there had been 18 subjects with 9 people in each group and with the same 6 scores as before (each score occurring three times)?

Application

Example (Dr. Koether's Smart Pill)

- What if there had been 18 subjects with 9 people in each group and with the same 6 scores as before (each score occurring three times)?
- In this case, there are 48,620 ways to assign 9 of the 18 subjects to the treatment group.

Application

Example (Dr. Koether's Smart Pill)

- What if there had been 18 subjects with 9 people in each group and with the same 6 scores as before (each score occurring three times)?
- In this case, there are 48,620 ways to assign 9 of the 18 subjects to the treatment group.
- Of the 48,620 ways, only 534 of them will result in a difference of at least 5 points.

Application

Example (Dr. Koether's Smart Pill)

- We should conclude that for samples of much larger sizes, there is virtually no chance that the treatment group would outperform the control group by at least 5 points, if Dr. Koether's Smart Pill really has no effect.

Application

Example (Dr. Koether's Smart Pill)

- We should conclude that for samples of much larger sizes, there is virtually no chance that the treatment group would outperform the control group by at least 5 points, if Dr. Koether's Smart Pill really has no effect.
- Therefore, if it actually did happen for large samples, we could be confident that the difference is due to Dr. Koether's Smart Pill, not chance.

Statistically Significant

Definition (Statistically Significant)

If an effect is so large that it is extremely unlikely that it occurred by chance, then we say that it is statistically significant.

Outline

Example

Example (IQ Scores)

- In the general population, IQ scores have a normal distribution with a mean of 100 and a standard deviation of 15.

Example

Example (IQ Scores)

- In the general population, IQ scores have a normal distribution with a mean of 100 and a standard deviation of 15.
- Suppose that in a specific community, high levels of lead were found in the drinking water.

Example

Example (IQ Scores)

- In the general population, IQ scores have a normal distribution with a mean of 100 and a standard deviation of 15.
- Suppose that in a specific community, high levels of lead were found in the drinking water.
- A sample of 500 people from that community is selected and their IQs are measured.

Example

Example (IQ Scores)

- In the general population, IQ scores have a normal distribution with a mean of 100 and a standard deviation of 15.
- Suppose that in a specific community, high levels of lead were found in the drinking water.
- A sample of 500 people from that community is selected and their IQs are measured.
- The researchers find an average IQ of 98.

Example

Example (IQ Scores)

- In the general population, IQ scores have a normal distribution with a mean of 100 and a standard deviation of 15.
- Suppose that in a specific community, high levels of lead were found in the drinking water.
- A sample of 500 people from that community is selected and their IQs are measured.
- The researchers find an average IQ of 98.
- Can they conclude with confidence that the IQ in that community is less than 100?

Example

Example (IQ Scores)

- In the general population, IQ scores have a normal distribution with a mean of 100 and a standard deviation of 15.
- Suppose that in a specific community, high levels of lead were found in the drinking water.
- A sample of 500 people from that community is selected and their IQs are measured.
- The researchers find an average IQ of 98.
- Can they conclude with confidence that the IQ in that community is less than 100?
- Can they conclude that the lower average IQ is because of the lead in the water?

Example

Example (IQ Scores)

- On a fair die, the numbers 1, 2, 3, 4, 5, and 6 turn up with equal frequency, in the long run.

Example

Example (IQ Scores)

- On a fair die, the numbers 1, 2, 3, 4, 5, and 6 turn up with equal frequency, in the long run.
- It turns out that the mean is 3.5 and the standard deviation is 1.708.

Example

Example (IQ Scores)

- On a fair die, the numbers 1, 2, 3, 4, 5, and 6 turn up with equal frequency, in the long run.
- It turns out that the mean is 3.5 and the standard deviation is 1.708.
- Suppose that we suspect that a particular die is not fair, but that it tends to show larger numbers (4, 5, 6) more often than it should.

Example

Example (IQ Scores)

- On a fair die, the numbers 1, 2, 3, 4, 5, and 6 turn up with equal frequency, in the long run.
- It turns out that the mean is 3.5 and the standard deviation is 1.708.
- Suppose that we suspect that a particular die is not fair, but that it tends to show larger numbers (4, 5, 6) more often than it should.
- We roll the die 40 times and find $\overline{x} = 3.9$.

Example

Example (IQ Scores)

- On a fair die, the numbers 1, 2, 3, 4, 5, and 6 turn up with equal frequency, in the long run.
- It turns out that the mean is 3.5 and the standard deviation is 1.708.
- Suppose that we suspect that a particular die is not fair, but that it tends to show larger numbers (4, 5, 6) more often than it should.
- We roll the die 40 times and find $\overline{x} = 3.9$.
- Can we conclude with confidence that the die is not fair?

Example

Example (IQ Scores)

- On a fair die, the numbers 1, 2, 3, 4, 5, and 6 turn up with equal frequency, in the long run.
- It turns out that the mean is 3.5 and the standard deviation is 1.708.
- Suppose that we suspect that a particular die is not fair, but that it tends to show larger numbers (4, 5, 6) more often than it should.
- We roll the die 40 times and find $\overline{x} = 3.9$.
- Can we conclude with confidence that the die is not fair?
- What if we rolled the die 100 more times and still obtained an average of 3.9?

Outline

Assignment

Assignment

- Read Sections 15.6.
- Apply Your Knowledge: 14.
- Exercises 36, 37, 38, 39, 40, 41.