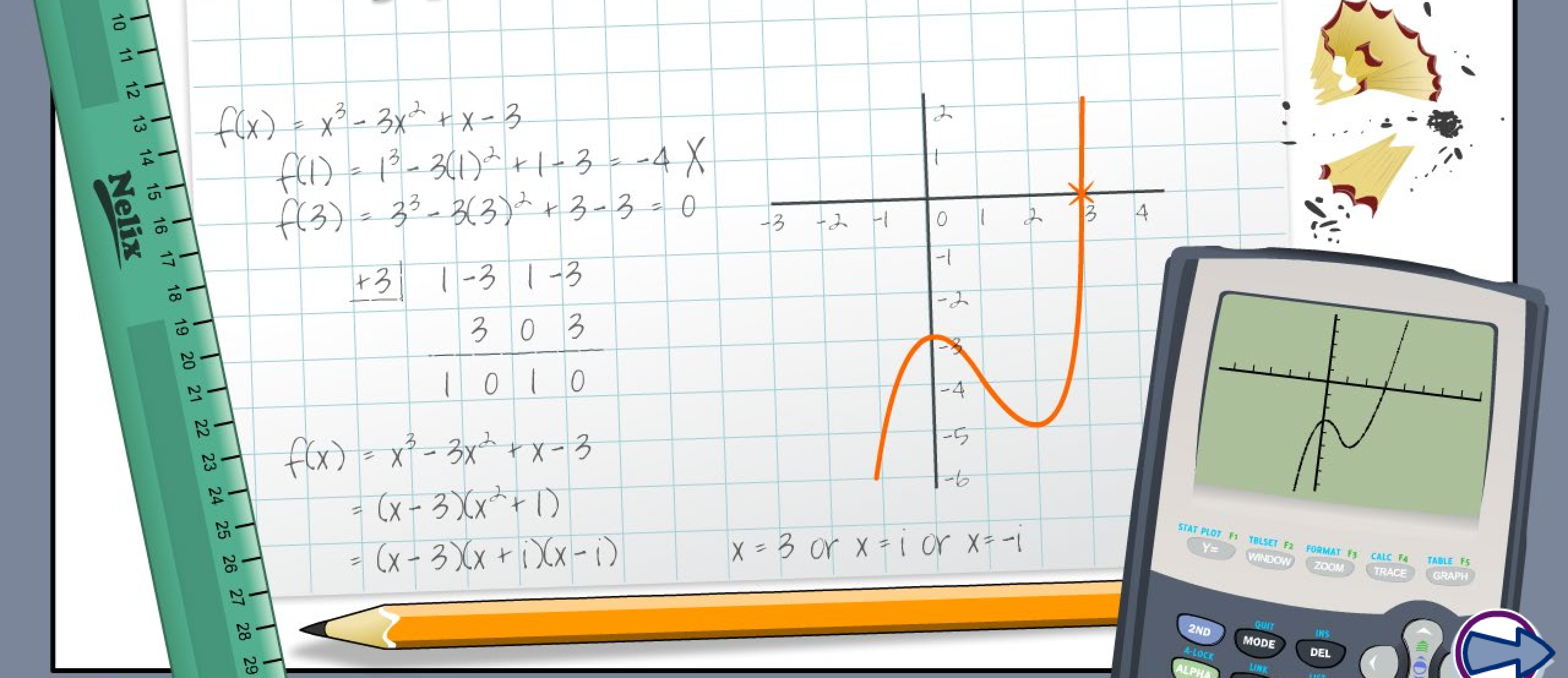
Hypothesis Testing



Common core icons

This icon indicates a slide where the Standards for Mathematical Practice are being developed. Details of these are given in the Notes field.

Slides containing examples of mathematical modeling are marked with this stamp.

This icon indicates an opportunity for discussion or group work.

The **Standards for Mathematical Practice** outlined in the Common Core State Standards for Mathematics describe varieties of expertise that mathematics educators at all levels should seek to develop in their students.

These are:

1) Make sense of problems and persevere in solving them.

2) Reason abstractly and quantitatively.

3) Construct viable arguments and critique the reasoning of others.

4) Model with mathematics.

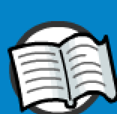
5) Use appropriate tools strategically.

6) Attend to precision.

7) Look for and make use of structure.

8) Look for and express regularity in repeated reasoning.

This icon indicates that the slide contains activities created in Flash. These activities are not editable.



This icon indicates teacher's notes in the Notes field.

Experiments

How can a company test how effective a new painkiller is?

They can design an experiment to rate the level of pain and administer the painkiller (a high pain level is the independent variable). They can then monitor the resulting pain level (the dependent variable).

A controlled experiment divides test subjects into two groups: the treatment group and the control group. There are several variables in an experiment:

- **control variables** – the researcher controls confounding or hidden variables so that they don't affect the outcome
- **independent variable** – the variable that the researcher manipulates
- **dependent** or **response variable** – the variable that responds to changes in the independent variable.

Principles of experimental design

In order for a experiment to provide the most accurate results, the experiment should have:

- **control** of sources of variation that may have an effect on the response variable, but that are not the factors being studied
- **randomization** of subjects and treatments to make up for the effects of unknown or uncontrollable sources of variation
- **replication** over as many subjects as possible
- **blocking** – some experiments organize test subjects into groups, or blocks, with similar features to further reduce variability.

Design an experiment

A company has 1,000 volunteers with pain scores of 9 or greater: 500 men and 500 women. Design an experiment without gender bias to test the hypothesis: "If the new painkiller is administered, then the pain score will decrease to 3 or less."

To remove the gender bias, the test subjects should be blocked and distributed as shown in the table.

	pain-killer	placebo
women	250	250
men	250	250

identify the independent variable:

whether they are given the painkiller

identify the response variable:

resulting pain score

If the subject has a pain score of 9 or over, they take the medicine. If their pain score decreases to 3 or less, the painkiller is considered successful.

Experimental design quiz

How well do you understand the principles of good experimental design?

Read the examples of experiments and identify the principle being used.

Press **start** to begin.

start

Designing experiments

For each of the following scenarios, design an experiment, identify the variables, and discuss possible problems that might occur in testing.

1) The effect of amount of sleep on stress level.

2) The effect of weather on a person's mood.

3) The effect of air pressure on plant growth.

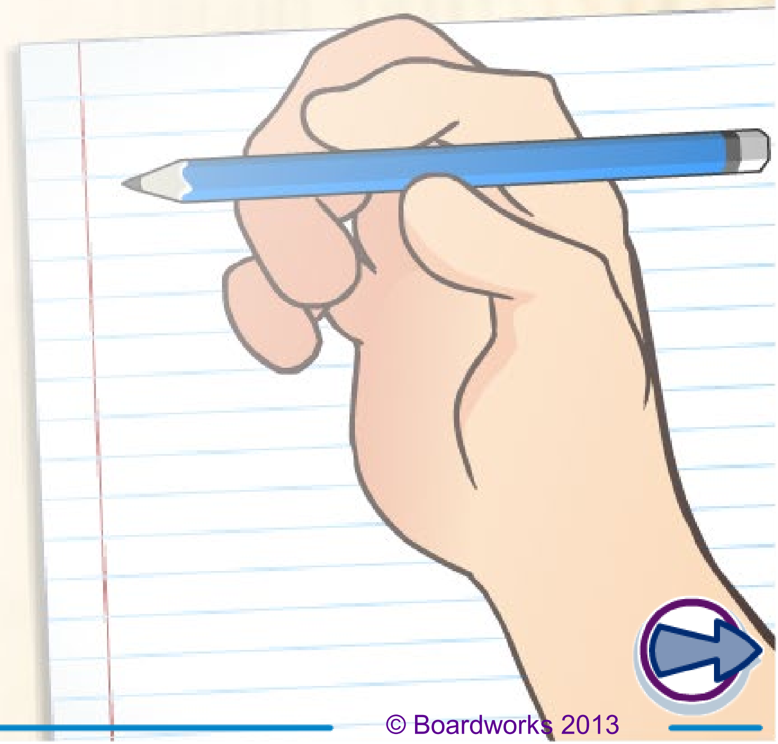
4) The effect of temperature in a classroom on test scores.

Testing a hypothesis

Experiments can be used to test a research hypothesis. A **statistical hypothesis** is an assumption about a population parameter that is verified based on the results of sample data.

Testing a statistical hypothesis involves three important steps:

- stating the hypothesis
- analyzing sample or experimental data
- interpreting results and deciding whether or not to reject the hypothesis.



Null and alternative hypotheses

In a statistical hypothesis test there are always two hypotheses.

null hypothesis (H_0):
states that there is no difference between the sample mean and the population mean.

alternative hypothesis (H_1):
states that there is a difference between the sample mean and the population mean.

The claim can be either the null or alternative hypothesis.

The statistical evidence can only support the claim if the claim is the alternative hypothesis, and it can only reject the claim if the claim is the null hypothesis.

Stating null and alternative hypotheses

A medical researcher wants to determine if a new medicine will raise, lower, or keep a patient's pulse rate the same. The researcher knows that the mean pulse rate for the population being studied is 79 beats per minute.

State the null and alternative hypothesis.

The null hypothesis states that the sample mean is equal to the population mean. The population mean pulse rate is equal to 79.

$$H_0: \mu = 79$$

For H_1 set the mean not equal to 79, to indicate any change in the mean pulse rate:

$$H_1: \mu \neq 79$$

Statistical tests

two-tailed tests | left-tailed tests | right-tailed tests

In science, research hypotheses are tested using experiments. In statistics, **statistical hypotheses** are tested using different types of statistical hypothesis tests. These tests help the statistician to decide whether to accept or reject their hypothesis.

Press the tabs to find out more about types of tests used to test statistical hypotheses.

Null and alternative hypotheses

For each scenario, state the null and alternative hypothesis.

Press the gray boxes to reveal the null and alternative hypotheses.

1) The average age of community school students is 23.5.

H_0 : ? H_1 : ?

2) The average age of teachers is greater than 29 years of age.

H_0 : ? H_1 : ?

3) The average cost of a used mp3 player is less than $75.

H_0 : ? H_1 : ?

Fertilizer and plant growth

The average rate of plant growth is 0.125 in/day. A researcher wants to test the hypothesis: "As the concentration of fertilizer increases, the rate of plant growth increases."

Here are her results:

trials	20	20	20	20
fertilizer concentration (g/L)	0	5	10	15
average rate of growth (in/day)	0.125	0.185	0.210	0.211

Identify the variables in the experiment.

control variables: plant, soil, water, light, temperature

independent variables: concentration of fertilizer

dependent variables: growth rate of plant

Write statistical hypotheses based on the data.

H_0: $\mu = 0.125$ H_1: $\mu > 0.125$ This is the claim.

Statistical tests and test values

Using z-scores is way to test hypotheses from any of the statistical tests.

The data from n samples is converted to a **test value**, which is the z-score for the sample mean.

$$\textbf{\textit{test value}} = \frac{\text{(sample mean – population mean)}}{\text{(standard deviation of the population } / \sqrt{n}\text{)}}$$

The test value is compared to a **critical value**, which is the z-score of the edge of the critical region.

If the test value is in the critical region, the null hypothesis is rejected.

Example

The average weight of strawberries over the last 10 years was 25 g, with standard deviation 10 g. From a sample of 45 strawberries, the average weight this year is 30 g. Convert this problem into a statistical test to check if this year's crop is heavier using a critical value of 1.64.

determine the type of test: This is a right-tailed test with critical value 1.64.

state the hypotheses: H_0: $\mu = 25$ H_1: $\mu > 25$ This is the claim.

find test value: $$\frac{\text{sample mean} - \text{population mean}}{\text{population standard deviation} \div \sqrt{n}}$$

$$= \frac{30 - 25}{10 \div \sqrt{45}} = \mathbf{3.35}$$

The test value exceeds the critical value and is in the critical region, so the null hypothesis is rejected, meaning that this year's crop is heavier.

What happens if a sample size of 10 is used?

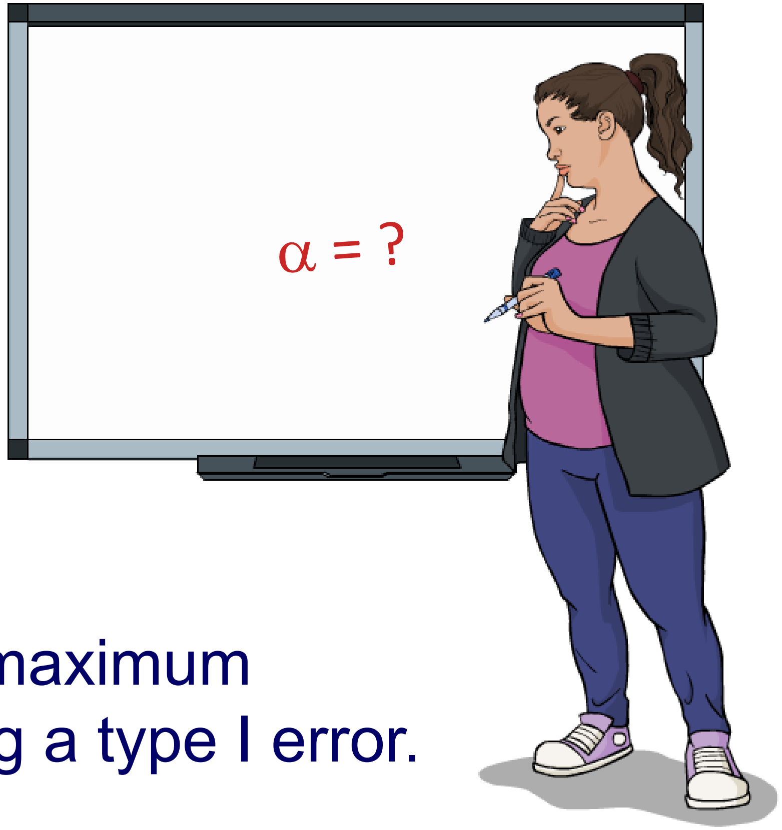
Errors

Errors can occur when testing a hypothesis:

- A **type I error** occurs if the null hypothesis is rejected when it is true. This is also known as a false positive.

- A **type II error** occurs if the null hypothesis is not rejected when it is false. This is also called a false negative.

- The **level of significance** is the maximum permitted probability of committing a type I error. This is symbolized by α.



False negative and positive

Experimental probability can be affected by outcomes known as **false negatives** and **false positives**.

A test for a pollen allergy is not 100% accurate.

- For people with the allergy it is 90% accurate, with 10% told that they are **not** allergic (false negatives).
- For people without the allergy it is 80% accurate, with 20% told that they **are** allergic (false positives).

Freddie tests positive for the allergy. What is the probability that he doesn't have the allergy, if 20% of the population actually has the allergy?

False positives

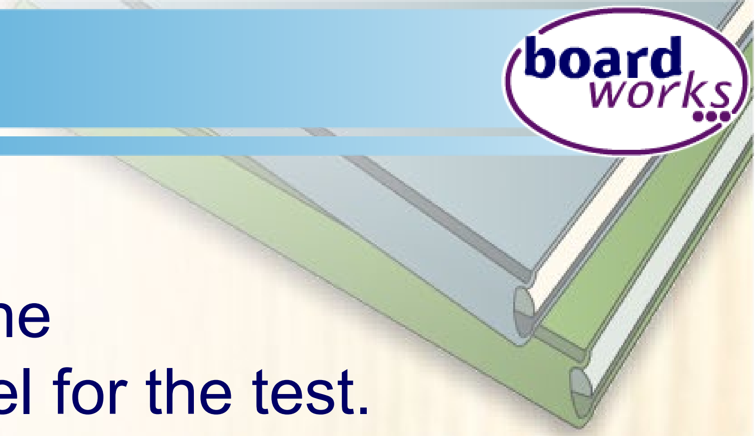
Fred tests positive for a pollen allergy. The test is 90% accurate for people with the allergy (10% receive a false negative), and 80% accurate for people without the allergy (20% receive a false positive).

If 20% of the population actually has the allergy, what is the chance that Fred's test is a false positive?

Press **start** to begin.

start

Significance level



After deciding on the type of test (right-tailed, left-tailed, or two-tailed) the statistician chooses a significance level for the test.

If there are cost or health concerns, a statistician may choose a smaller significance level to reduce the chance of error.

Statisticians generally use three different levels of significance:

- 0.10
- 0.05
- 0.01

The level of significance is used to define the critical region.

Critical values

critical value of a **one**-tailed test

critical values of a **two**-tailed test

Press on a yellow button to see how to calculate the C.V. for each type of test.

Once a significance level has been decided, the researcher can find the **critical value (C.V.)** for the appropriate statistical test.

The critical value will determine the critical and non-critical region under the normal distribution.

Statistical test examples

Question 1/2:

A sample of 30 police officers has a mean salary of $52,000. The standard deviation of the population is $5,140.
At α = 0.05, test the claim that police officers earn more than $52,500 per year.

Press the "**=**" button to show
the calculations step-by-step.

Knee rehabilitation

It has been said that the average cost of rehabilitation for a knee surgery patient is \$11,265. The standard deviation of the population is \$3,250. To see if the average cost of rehabilitation is different at a particular facility, a researcher selects a random sample of 35 knee surgery patients and finds that the average cost of their rehabilitation is \$12,500.

At $\alpha = 0.01$, can it be concluded that the average cost of knee rehabilitation at a particular facility is different from \$11,265?

Press the "next" arrow to see the solution step-by-step.