More Tests for Averages: Comparing Two Samples

Where are we going?

Review of concepts of hypothesis testing

Comparing two independent samples

- SE of a difference
- Applications to percentages and continuous measurements

Randomized experiments

Review

A **significance test** is aimed at determining whether a result is real or could possibly be due to chance.

The **null hypothesis** says that the result is due to chance. A probability calculation is made under the null hypothesis via a box model.

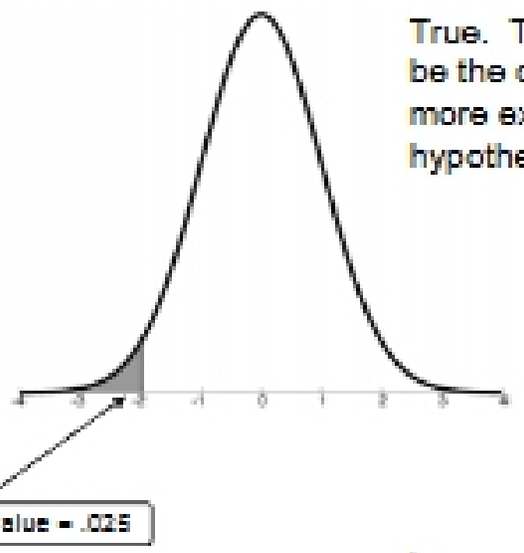
A **test statistic** measures the difference between the data and what would be expected under the null hypothesis.

The **observed significance level**, or **P-value**, is the chance of getting a test statistic as or more extreme than the one observed if the null hypothesis were true.

Small P-values are evidence against the null hypothesis.

Sometimes a threshold (like 5%) is set in advance, and the null hypothesis is said to be **rejected** if the P-value is smaller than that threshold.

Question: If the null hypothesis is true, the chance of getting a P-value of .025 or smaller is .025. True or False?



True. The P-value is *defined* to be the chance of getting as or more extreme a result if the null hypothesis is true.

11. (Hard.) Discount stores often introduce new merchandise at a special low price in order to induce people to try it. However, a psychologist predicted that this practice would actually reduce sales. With the cooperation of a discount chain, an experiment was performed to test the prediction.* Twenty-five pairs of stores were selected, matched according to such characteristics as location and sales volume. These stores did not advertise, and displayed their merchandise in similar ways.

A new kind of cookie was introduced in all 50 stores. For each pair of stores, one was chosen at random to introduce the cookies at a special low price, the price increasing to its regular level after two weeks; the other store in the pair introduced the cookies at the regular price. Total sales of the cookies were computed for each store for six weeks from the time they were introduced.

In 18 of the 25 pairs, the store which introduced the cookies at the regular price turned out to have sold more of them than the other store. Can this result be explained as a chance variation? Or does it support the prediction that introducing merchandise at a low price reduces long-run sales? (Formulate the null hypothesis as a box model; there is no alternative hypothesis about the box.)

Summary: 25 pairs of stores. One member (chosen randomly dropped the price). In 18 cases of 25, the store that did not drop the price sold more.

Null hypothesis: the difference is due to chance

Alternative hypothesis: there is some kind of effect.

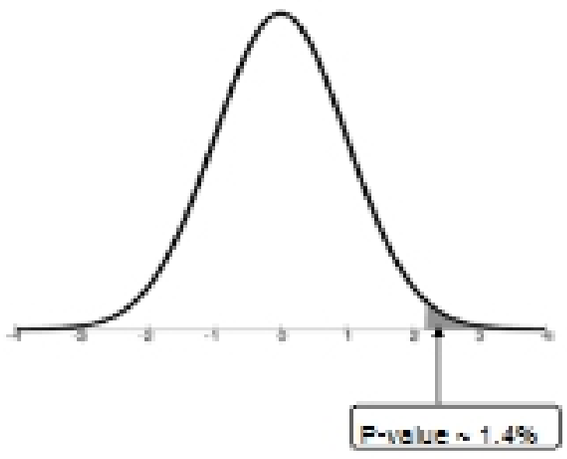
Probability model for the null hypothesis: 25 draws with replacement from what kind of box?

Calculations under the assumption that the null hypothesis is true: what can we say about the number of times a store that did not drop the price sold more?

EV of # =

SE of # = 1

Test statistic:



Conclusion: such a result is quite unlikely if there is no effect. There is thus strong evidence against the null hypothesis.