How can we tell whether to reject or not reject the null hypothesis?

When the null hypothesis is true, a certain set of values occur for the test statistic with a specific probability structure. In a Z hypothesis test this set of test statistic values that is expected when the null hypothesis is true is the Z-distribution, also called the Standard Normal distribution or N(0,1) to signify a normal curve with mean of 0 and variance of 1.

If the null hypothesis, Ho, is false
then the observed test statistic value will be
in an unlikely region of the Z distribution.

REJECT Ho when the test statistic is in an unlikely region in the Z distribution.

If the null hypothesis, Ho, is true
then the observed Z test statistic will be
in a likely region of the Z distribution.

DO NOT REJECT Ho when the test statistic is in a likely region in the Z-distribution.

Rejection Region Process:

Where are the unlikely regions of the Z distribution?

The unlikely regions are the tail area(s) but which tail(s) and how big depend on two things. In a two-tail test both sides of the distribution are involved, in a right-tail test only the right side, and in a left-tail test only the left side. How big depends on how low the level of unlikeliness is set, which is the significance level of the test, alpha or in a symbol, α.

How is the significance level, α, set?

What is the maximum risk of rejecting a true null hypothesis that can be tolerated considering the consequences of claiming Ho false and Ha true when that was not the case? That is the significance level of a hypothesis test. It is called the alpha-value.

How are Rejection Regions set up and used?

Identify the critical values that cut off the area equal to α in the tail determined by the direction of the test, that is whether it is two-, right-, or left-tail. If the test statistic is in the unlikely region identified by the critical value then the null hypothesis is rejected at the associated significance level.

Reject the null hypothesis, Ho, if $P<\alpha$, but what is the P-value of the test?

P-Value Process:

What is the likelihood of seeing a test statistic at least as extreme as the one actually observed if the null hypothesis is true?

This observed likelihood of data at least as extreme as the data observed, given the null hypothesis is true, is called the p-value of the hypothesis test. The smaller the p-value the stronger the evidence against the null hypothesis.

How do you calculate the p-value of a hypothesis test?

The p-value of a hypothesis test is the tail area(s) associated with the test statistic. In a two-tail test the area above the positive value of the test statistic and the area below the negative value of the test statistic are added together to comprise the p-value. In a left-tail test the p-value is the area to the left of the test statistic. In a right-tail test the p-value is the area to the right of the test statistic.

How do you use the p-value to make the rejection decision?

The p-value is the chance you are wrong if you decide to reject the null hypothesis based on the data observed. Can you tolerate this chance of error when rejecting the null hypothesis? If this level of risk of error can be tolerated then the null hypothesis can be rejected.

How do you know when the p-value represents a reasonable chance of error and that the null hypothesis should be rejected?

What are the consequences of being wrong if you claim that the null hypothesis is false and the alternative is true, but that is not the case? If those consequences are extreme the p-value at which the null hypothesis is rejected should be very small.

The smaller the p-value the more evidence against the null hypothesis.