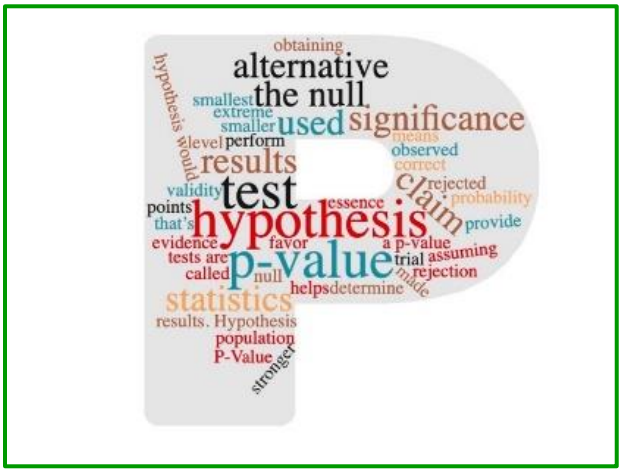
P-VALUES SIMPLIFIED



Preface

This discussion of p-values is viewed in the context of a hypothesis test. While the concepts can be extended to more complex models, this simple test is employed for the sake of clarity and ease of comprehension.

In the realm of hypothesis testing one can never prove that a null hypothesis is true; rather, we conduct a test under the assumption that the null hypothesis is correct and endeavor to see if there is an acceptable level of probability to reject it.

Strictly speaking, if we don't find evidence to reject, we cannot say that we "accept the null". The aim of hypothesis testing is <u>not</u> to verify the null hypothesis; but, rather to examine (test) some claim/condition delineated via an alternative hypothesis.

In the following discussion we sometimes refer to the null hypothesis as being true, when in fact, we do not know and are not testing if it is true.

Quick Definitions

The term p-value is an abbreviation for Probability Value.
Hence, the concept of a p-value concerns the notion of probability.

Probability is defined as:

The branch of mathematics concerning events and numerical descriptions of how likely they are to occur.
The probability of an event is a number between 0 and 1 (inclusive).
The larger the probability, the more likely an event is to occur.

In mathematical terms, the probability value domain is -- $[\ 0 \leq \text{p-value} \leq 1\ ]$

FORMAL DEFINITIONS

The formal definition often proffered defines a p-value as:

The probability of obtaining test results at least as extreme as the result actually observed, under the assumption that the null hypothesis is correct.

Similarly, it is often postulated:

***We are assuming the null hypothesis is true and trying to see if there is evidence against it*.**

Confused yet?... Hang it there!

This "evidence against it" wording could be more precisely delineated as "evidence of statistical significance against the null".

Following this line of reasoning, smaller p-values provide greater evidence for rejecting the null hypothesis.

However, the idea (and wording) of a smaller p-value providing greater evidence against the null hypothesis is counterintuitive.

For many people not versed is the field of statistics, such formal definitions may contribute to high levels of confusion and frustration.

A MORE UNDERSTANDABLE DEFINITION

A clearer and more intuitive definition of a p-value answers the question: Assuming the null hypothesis is true, what do the sample data say about how likely the null hypothesis is to be true?

When viewed in the context of this question, a p-value is more simply defined as:

THE PROBABILITY TO WHICH THE DATA SUPPORT THE NULL HYPOTHESIS

Accordingly, a large p-value lends support to the assertion of a correct null hypothesis. Hence, larger p-values result in failure to reject the null hypothesis.

****

Conversely, a small p-value means that there is a lesser chance that the data support the null hypothesis. Thereby lending acceptance of the alternative hypothesis.

Remember, a p-value doesn't tell you if the null hypothesis is true or false. It just tells you how likely it would be to obtain a particular result (from sample data) <u>if the null hypothesis were true</u>.
A p-value is a piece of evidence, not a definitive proof.

Examples to Help Solidify this Discussion.

- **A result of p = 0.836 shows an 86.3% probability that the data support the null. This is strong evidence supporting the null hypothesis. (And weak evidence to reject it).**

- **A result of p = 0.022 shows a 2.2% probability that the data support the null. This is weak evidence supporting the null. Accordingly, this would most likely lead to rejecting the null and accepting the alternative hypothesis.**