Comments on Hypothesis Testing

The Steps

Hypothesis testing always follows the same steps. Sometimes some of the steps are either skipped or are implicit. But that can easily lead to confusion and misunderstanding. For clarity it is important to state all steps explicitly.
The steps are as follows:

- **A *statistical model* is stated. Statistical models involve *random variables* and non-random *parameters*.**
- **Two hypotheses are stated; a *null hypothesis* H_0 and an *alternative hypothesis* H_1. A hypothesis is simply a *statement about the model parameters*. The two hypotheses should be mutually exclusive, do not overlap (to avoid interpretational and technical problems).**
- **A type-I error is chosen in advance; before observing any data (before "seeing the data").**
- **A *test statistic* is computed.**

- **The distribution of the test statistic under the null hypothesis must be known (exactly or approximately). That distribution is used to decide whether to accept H_0 and reject H_1, or accept H_1 and reject H_0 (there are no other options).**

Significance Testing

The above description was of the Neyman-Pearson formulation of hypothesis testing as a decision problem - to choose either H_0 or H_1. In that approach there are two hypotheses and two types of error (type-I and type-II).
Another, historically older, approach is Fisher's *significance testing*. In that approach there is only one hypothesis, H_0, and only one type of error, type-I. The result is to either *reject* H_0 or *fail to reject* H_0. A *p-value* can be computed as well (based on the distribution of the test statistic under the null hypothesis).
The above distinction explains the seemingly conflicting messages students receive in introductory classes. One message is "you can't accept H_0, you either reject or fail to reject H_0". The other message is "you can reject H_0 and accept H_1, or reject H_1 and accept H_0". The first message is based on Fisher's approach while the second is based on the Neyman-Pearson approach. A common source of confusion is mixing the two types of hypothesis testing.