Hypothesis Testing

We begin with a statement – the null hypothesis – "on trial." At the end of the trial, we will find that the evidence at hand either contradicts the statement to some extent (i.e., the evidence supports a finding of "guilty"), or doesn't really contradict the statement (i.e., the evidence doesn't support a finding of "guilty," so we find it "not (shown to be) guilty").

Think of a hypothetical world in which the statement on trial is true. (If there's more than one such world, choose the one which most closely fits the observed data.)

The *significance level* of the data (with respect to the statement on trial) is

$$\text{Prob}\left(\begin{array}{c}\text{we'd see data at}\\ \text{least as contradictory}\\ \text{to the statement as is}\\ \text{the data at hand}\end{array}\;\middle|\;\begin{array}{c}\text{the study that yielded the data}\\ \text{at hand were to be conducted}\\ \text{in the hypothetical world}\\ \text{where the statement is true}\end{array}\right)$$

We interpret the significance level of the data using this "translation" table:

If the numeric significance level of the data is	**then the data, all by itself, makes us**	**and the data supports the alternative**
above 20%	not at all suspicious	not at all
between 10% and 20%	a little bit suspicious	a little bit
between 5% and 10%	moderately suspicious	moderately
between 2% and 5%	very suspicious	strongly
between 1% and 2%	extremely suspicious	very strongly
below 1%	overwhelmingly suspicious	overwhelmingly

We never conclude that the evidence *supports* the statement on trial (i.e., the statement is never found "innocent"). Therefore, if our ultimate goal is to see if evidence supports a statement, we must put the opposite statement on trial, and see if the evidence contradicts that opposite statement.