TOOL

Evaluating the Consequences of an Evidence-Based Decision

Often, data scientists need to decide between two courses of action based on data, and the decisions they make can have important impacts on the real world. Since these decisions are made about a larger population from a smaller sample, there is always a possibility of making an error. But different types of errors have different consequences.

For example, consider a randomized control trial (RCT). You have recruited a few patients, found that the new drug seems to work somewhat better than the old one in your sample. Perhaps you even calculated the p-value and found that it is somewhat small — say, 10%. At this point, you need to choose between two decisions:

> **Null Hypothesis (H_o):** The new treatment does not work better than the old treatment in the population.
>
> **Alternative Hypothesis (H_A):** The new treatment does work better than the old treatment in the population.

If you decide to select the H_A, you would probably recommend the new treatment for the next phase of trial. If you select the H_o, you would probably suggest conducting another trial with larger sample size, or abandon further research on this treatment.

Depending on which of the two hypotheses is true and which of the two decisions you select, there are four possible outcomes. The **Truth Table** below contains all four of these possibilities:

Truth Table

		Our Decision: Null Hypothesis	Our Decision: Alternative Hypothesis
The Truth	Null Hypothesis	Correct	False positive (Type I)
The Truth	Alternative Hypothesis	False negative (Type II)	Correct

1. The top-left quadrant of the Truth table applies when you select the H_o when the H_o is true. This scenario is called a **true negative (TN)**. This means your decision is correct, but does not change your current understanding of the situation (status quo). In the randomized control trial (RCT) example, this aligns with the new treatment working as well as the old treatment.
2. The top-right quadrant of the table applies when you select the H_A but in reality the H_o is correct. This is called a **false positive (FP)** or type-I error, and this type of error can lead to wasting time and money. In the RCT example, this would mean that you concluded the new drug works better than the old drug, when in reality it does not work better. This conclusion would lead you to continue testing, which would be a waste of resources since the drug provides no additional benefit.
3. The bottom-left quadrant of the table occurs when you select the H_o but H_A is correct. This is called a **false negative (FN)** or type-II error, and means you miss out on making a new discovery. In the RCT example, this type of error would occur if you conclude the new drug does not work better when it actually does work better than the old drug. You'd miss out on improving lives and possibly on increasing revenue.
4. The bottom-right quadrant of the table applies when you accurately reject the H_o in favor of the H_A. For example, the new drug does work better than the old one, and that is what you conclude from your data. This scenario is called a **true positive (TP)**.

In practice, data scientists carefully study the chances of the two types of error when making data driven decisions. For example, when testing hypotheses, one rejects the H_o in favor of the H_A if the p-value is less than a cut-off value. This cut-off value is chosen beforehand to control the chances of FP (type-I error) at a desired level (e.g. 1% or 5%). Data scientists also conduct statistical power analysis to examine the effect of the signal strength and sample size on the chances of FN (type-II error).