

1

Evidence and Verdicts

Before we focus in on using statistics as evidence to be used in making judgments, let's take a look at a widely used "verdict outcomes framework." This general framework is useful for framing judgments in a wide range of situations, including those encountered in statistical analysis.

Anytime we use evidence to arrive at a judgment, there are four generic outcomes possible, as shown in Table 1.1. Two outcomes correspond to correct judgments and two correspond to incorrect judgments, although we rarely know whether our judgments are correct or incorrect. Consider a jury trial in U.S. criminal court. Ideally, the jury is always correct, judging innocent defendants not guilty and judging guilty defendants guilty. Evidence is never perfect, though, and so juries will make erroneous judgments, judging innocent defendants guilty or guilty defendants not guilty.

Table 1.1 Verdict outcomes framework.

The verdict Unknown Truth	Not guilty verdict (do not reject presumption)	Guilty verdict (reject presumption)
Defendant is innocent (the presumption)	Correct	Incorrect verdict (type I error)
Defendant is guilty	Incorrect verdict (type II error)	Correct

In U.S. criminal court, the presumption is that a defendant is innocent until "proven" guilty. Further, convention in U.S. criminal court has it that we are more afraid of punishing an innocent person (type I error) than we are of letting a guilty person go unpunished (type II error). Because of this fear, the threshold for a guilty verdict is set high: "Beyond a reasonable doubt." So, convicting an

innocent person should be a relatively unlikely outcome. In U.S. criminal court, we are willing to have a greater chance of letting a guilty person go unpunished than we are of punishing an innocent person. In short, we need to be very sure before we reject the presumption of innocence and render a verdict of guilty in U.S. criminal court.

We can change the relative chances of the two types of error by changing the threshold. Say we change from “beyond a reasonable doubt” to “a preponderance of evidence.” The former is the threshold used in U.S. criminal court, and the latter is the threshold used in U.S. civil court. Let’s say that the former corresponds to being 95% sure before judging a defendant guilty and that the latter corresponds to being 51% sure before judging a defendant guilty. You can imagine cases where the same evidence results in different verdicts in criminal and civil court, which indeed does happen. For example, say that the evidence leads to the jury being 60% sure of the defendant’s guilt. The jury verdict in criminal court would be not guilty ($60\% < 95\%$) but the jury verdict in civil court would be guilty ($60\% > 51\%$). Compared to criminal court, civil court is more likely to declare an innocent person guilty (type I error), but is also less likely to declare a guilty person not guilty (type II error).

Changing the verdict threshold
either decreases type I error while increasing type II error (criminal court)
or increases type I error while decreasing type II error (civil court)

Statistical analysis is conducted as if in criminal court. Below are a number of jury guidelines that have parallels in statistical analysis, as we’ll see repeatedly.

Erroneously rejecting the presumption of innocence (type I error) is feared most. The possible verdicts are “guilty” and “not guilty.” There is no verdict of “innocent.” Reasons other than perceived innocence can lead to a not guilty verdict, such as insufficient evidence.

To reject the presumption of innocence and render a guilty verdict, there must be a sufficient amount of (unbiased) evidence.

To reject the presumption of innocence and render a guilty verdict, the pieces of evidence must be sufficiently consistent (not at variance with each other).

Statistical analysis formally evaluates evidence in order to determine whether to reject or not reject a stated presumption, and it is primarily concerned with limiting the chances of type I error. Further, the amount of evidence and the variance of evidence are key characteristics of evidence that are formally incorporated into the evaluation process. In what follows, we’ll see how this is accomplished.