A Primer on Probability for Design Inferences

By William A. Dembski

[version 1.1]

Distinction between Outcomes and Events

Probabilities are numbers between 0 and 1 that attach to events. Events always occur with respect to a reference class of possibilities. Consider a die with faces 1 through 6. The reference class of possibilities in this case can be represented by the set {1, 2, 3, 4, 5, 6}. Any subset of this reference class then represents an event. For instance, the event E_{odd}, i.e., "an odd number was tossed," corresponds to {1, 3, 5}. Such an event is said occur if any one of its outcomes occurs, i.e., if either a 1 is tossed or a 3 or a 5. Outcomes can therefore be represented as singleton sets, i.e., sets with only one element. Thus, the outcomes associated with $E_{odd} = \{1, 3, 5\}$ are $E_1 = \{1\}$, $E_3 = \{3\}$, and $E_5 = \{5\}$. Outcomes are sometimes also called elementary events. Events include not only outcomes but also composite events like E_{odd} that include more than one outcome.

The Axioms of Probability

Probabilities obey the following axioms: (1) The impossible event (i.e., an event that entails a physical or logical impossibility) is represented by the empty set and has probability zero. (2) The necessary event (i.e., an event that is guaranteed to happen) is represented by the entire reference class of possibilities and has probability one (e.g., with the die example, $E_{nec} = \{1, 2, 3, 4, 5, 6\}$ has probability one). Events that are mutually exclusive have probabilities that sum together. Thus, in the previous example, $\mathbf{P}(E_{odd}) = \mathbf{P}(E_1) + \mathbf{P}(E_3) + \mathbf{P}(E_5)$ (i.e., $\mathbf{P}(\{1, 3, 5\}) = \mathbf{P}(\{1\}) + \mathbf{P}(\{3\}) + \mathbf{P}(\{5\})$). Important: mutually exclusive and exhaustive events always sum to one.

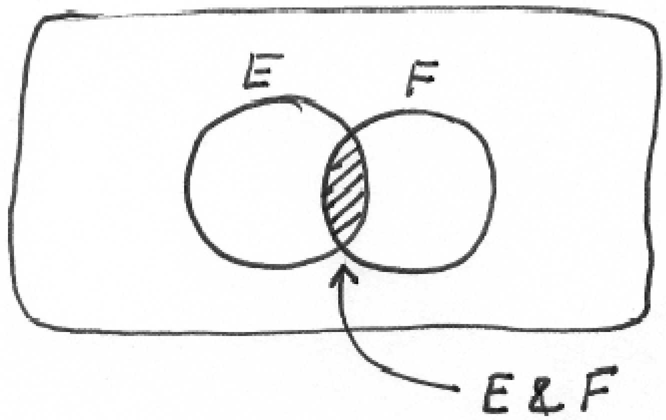
Interpretation of Probability

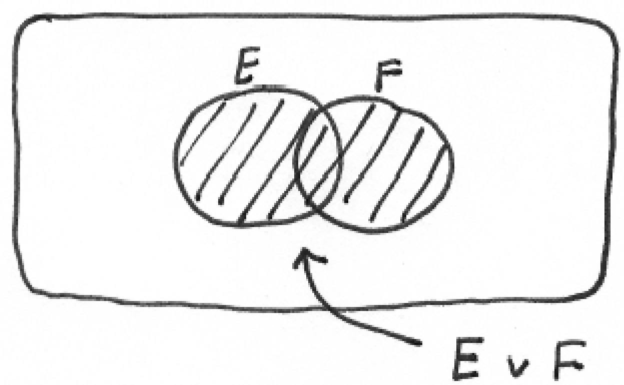
Probabilities are interpreted in three principal ways: (1) Degree of belief -- probability measures strength of belief that an event will occur. (2) Frequentist approach -- probability is a relative frequency (i.e., the number of occurrences of an event divided by the number of observed oppporutnities for the event to occur; relative frequencies are also called empirical probabilities). (3) Theoretical approach -- probability derives from properties of the system generating the events (e.g., dies are rigid, homogeneous cubes whose symmetry confers probability 1/6 on each face; quantum mechanical systems have probabilities derived from eigenvalues associated with the eigenstates of an observable).

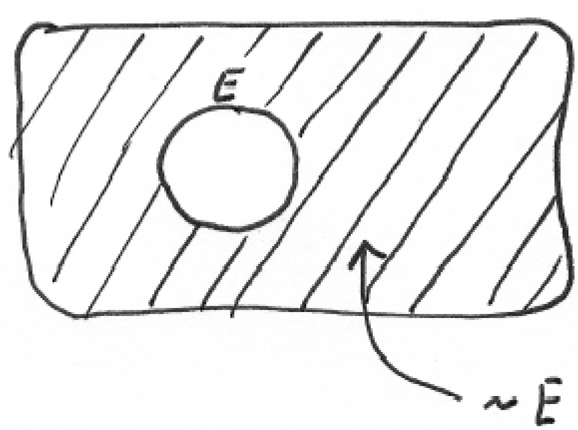
Conjunction, Disjunction, and Negation

Events can be modified by conjunction, disjunction, and negation. E & F is the conjunction (or intersection, also written $E \cap F$) of E and F and denotes the event such that both E and F occur.

$E \vee F$ is the disjunction (or union, also written $E \cup F$) of E and F and denotes the event such that either E or F or both occur. $\sim E$ is the negation (or complement, also written E^c) of E and denotes the the event that excludes E's occurrence. In pictures:





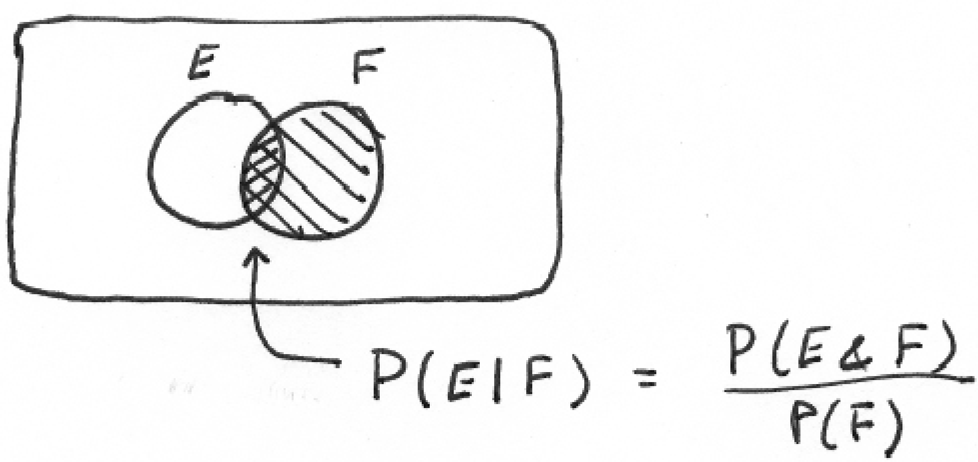


Conditional Probability

Suppose event F is known to have occurred and we then ask what is the probability of E. In that case, the reference class of possibilities contracts to F, and the probability of E is no longer simply $\mathbf{P}(E)$ (i.e., the probability of E within the original reference class), but the probability of E within the new reference class F, called the conditional probability of E given F and written $\mathbf{P}(E|F)$. This probability is by definition

$$\mathbf{P}(E \mid F) = \frac{\mathbf{P}(E \,\&\, F)}{\mathbf{P}(F)}$$

Graphically, this probability can be represented as follows:



Probabilistic Independence

As we have seen, for mutually exclusive events probabilities add. Specifically, the probability of a disjunction is the sum of the probabilities of the disjuncts. Thus, if $E_1, E_2, \ldots, E_n$ are mutually exclusive, $\mathbf{P}(E_1 \vee E_2 \vee \cdots \vee E_n) = \mathbf{P}(E_1) + \mathbf{P}(E_2) + \cdots + \mathbf{P}(E_n)$. Does a corresponding relationship hold for conjunction? For $E_1, E_2, \ldots, E_n$ arbitrary events such that no conjunction of them has zero probability, it follows from the definition of conditional probability that

$$\mathbf{P}(E_1 \,\&E_2 \,\&\cdots\,\&E_n) = \mathbf{P}(E_1) \times \mathbf{P}(E_2|\, E_1) \times \mathbf{P}(E_3|\, E_1 \,\&E_2) \times \cdots \times \mathbf{P}(E_n|\, E_1 \,\&E_2 \,\&\cdots\,\&E_{n-1}).$$

For instance, for just E_1 and E_2,

$$\begin{aligned}
\mathbf{P}(E_1 \,\&E_2) &= 1 \times \mathbf{P}\,(E_1 \,\&E_2) \\
&= [\mathbf{P}(E_1)/\,\mathbf{P}(E_1)] \times \mathbf{P}(E_1 \,\&E_2) \\
&= \mathbf{P}(E_1) \times [\mathbf{P}\,(E_1 \,\&E_2) \,/\, \mathbf{P}(E_1)] \\
&= \mathbf{P}(E_1) \times \mathbf{P}(E_2|\, E_1)
\end{aligned}$$

If, now, $\mathbf{P}(E_2| E_1) = \mathbf{P}(E_2)$, it follows that

$$\mathbf{P}(E_1 \& E_2) = \mathbf{P}(E_1) \times \mathbf{P}(E_2)$$

In that case, we say that E_1 and E_2 are probabilistically (or stochastically) independent. In general, we say that events $E_1, E_2, \ldots, E_n$ are independent if for all distinct events taken from this class, i.e., $E_{i_1}, E_{i_2}, \ldots, E_{i_k}$, $1 \leq k \leq n$,

$$\mathbf{P}(E_{i_1} \& E_{i_2} \& \cdots \& E_{i_k}) = \mathbf{P}(E_{i_1}) \times \mathbf{P}(E_{i_2}) \times \cdots \times \mathbf{P}(E_{i_k}).$$

Events are probabilistically independent if they derive from causally independent processes. The converse, however, is not be true -- events can be probabilistically independent without being causally independent.

Equiprobability and Uniform Probability

In many situations, individual outcomes (elementary events) each have the same probability. In that case, if there are N possible outcomes, each outcome has probability $1/N$. Equiprobability in this sense is a special case of uniform probability in which isomorphic events under some equivalence relation have identical probability (see my 1990 article on uniform probability at http://www.designinference.com/documents/2004.12.Uniform_Probability.pdf).

The Fisherian Approach to Design Inferences

This is the approach I adopt and have developed. In this approach there are always two events: an event E that the world presents to us and an event T that includes E (i.e., the occurrence of E entails the occurrence of T) and that we are able to identify via an independently given pattern (i.e., a pattern that we can reproduce without having witnessed E). Think of E as an arrow and T as a fixed target. If E lands in T and the probability of T is sufficiently small, i.e., $\mathbf{P}(T)$ is close to zero, then, on my approach, a design inference is warranted. For the details, see my article at http://www.designinference.com/documents/2005.06.Specification.pdf titled "Specification: The Pattern That Signifies Intelligence."

Bayes's Theorem

Given an event E and chance hypotheses $\mathbf{H}_1, \mathbf{H}_2, \ldots, \mathbf{H}_n$ that are mutually exclusive and exhaustive, the probability of any one of these hypotheses $\mathbf{H}_i$ given E is given by

$$\mathbf{P}(\mathbf{H}_i \mid E) = \frac{\mathbf{P}(E \mid \mathbf{H}_i)\mathbf{P}(\mathbf{H}_i)}{\mathbf{P}(E)}.$$

This is the simple form of Bayes's theorem. Since $\mathbf{H}_1, \mathbf{H}_2, \ldots, \mathbf{H}_n$ are mutually exclusive and exhaustive, it follows that the denominator here can be rewritten as

$$\begin{aligned} \mathbf{P}(E) &= \mathbf{P}([E \,\&\, \mathbf{H}_1] \vee [E \,\&\, \mathbf{H}_2] \vee \cdots \vee [E \,\&\, \mathbf{H}_n]) \\ &= \mathbf{P}(E \,\&\, \mathbf{H}_1) + \mathbf{P}(E \,\&\, \mathbf{H}_2) + \cdots + \mathbf{P}(E \,\&\, \mathbf{H}_n) \\ &= \mathbf{P}(E|\mathbf{H}_1)\mathbf{P}(\mathbf{H}_1) + \mathbf{P}(E|\mathbf{H}_2)\,\mathbf{P}(\mathbf{H}_2) + \cdots + \mathbf{P}(E|\mathbf{H}_n)\,\mathbf{P}(\mathbf{H}_n) \end{aligned}$$

These equalities follow simply from unpacking the axioms of probability and the definition of conditional probability. Substituting this last expression for the denominator in the simple form of Bayes's theorem now yields the standard form of Bayes's theorem:

$$\mathbf{P}(\mathbf{H}_i \mid E) = \frac{\mathbf{P}(E \mid \mathbf{H}_i)\mathbf{P}(\mathbf{H}_i)}{\mathbf{P}(E \mid \mathbf{H}_1)\mathbf{P}(\mathbf{H}_1) + \mathbf{P}(E \mid \mathbf{H}_2)\,\mathbf{P}(\mathbf{H}_2) + \cdots + \mathbf{P}(E \mid \mathbf{H}_n)\,\mathbf{P}(\mathbf{H}_n)}.$$

The Bayesian Approach to Design Inferences

In this approach, one considers an event E and hypotheses $\mathbf{H}_1$ and $\mathbf{H}_2$. Think of $\mathbf{H}_1$ as a design hypothesis and $\mathbf{H}_2$ as a chance/evolutionary hypothesis. Moreover, think of the event E as evidence for either of these hypotheses. To decide whether the evidence E better supports either $\mathbf{H}_1$ or $\mathbf{H}_2$ therefore amounts to comparing the probabilities $\mathbf{P}(\mathbf{H}_1|E)$ and $\mathbf{P}(\mathbf{H}_2|E)$ and determining which is bigger. These probabilities are known as *posterior probabilities* and measure the probability of a hypothesis given the event/evidence/data E.

Posterior probabilities cannot be calculated directly but must rather be calculated on the basis of Bayes's theorem. Using the simple form of Bayes's theorem, we find that the posterior probability $\mathbf{P}(\mathbf{H}_i|E)$ ($i = 1$ or 2) is expressed in terms of $\mathbf{P}(E|\mathbf{H}_i)$, known as the *likelihood* of $\mathbf{H}_i$ given E, and $\mathbf{P}(\mathbf{H}_i)$, known as the *prior probability* of $\mathbf{H}_i$. Often prior probabilities cannot be calculated directly. Moreover, in calculating the posterior probability, we still need to compute the denominator in the simple form of Bayes's theorem, namely, $\mathbf{P}(E)$.

Fortunately, this last term does not need to be calculated. Because the aim is to determine which of these hypotheses is better supported by the evidence, it is enough to form the ratio of posterior probabilities

$$\frac{\mathbf{P}(\mathbf{H}_1 \mid E)}{\mathbf{P}(\mathbf{H}_2 \mid E)}$$

and determine whether it is greater than or less than 1. If this ratio is greater than 1, it supports the hyothesis in the numerator ($\mathbf{H}_1$, which we are treating as the design hypothesis). If it is less than 1, it supports the hypothesis in the denominator ($\mathbf{H}_2$, which we are treating as the chance/evolution hypothesis).

This ratio, using the simple form of Bayes's theorem, can now be rewritten as follows (note that the denominator in Bayes's theorem simply cancels out):

$$\frac{\mathbf{P}(\mathbf{H}_1 \mid E)}{\mathbf{P}(\mathbf{H}_2 \mid E)} = \frac{\mathbf{P}(E \mid \mathbf{H}_1)}{\mathbf{P}(E \mid \mathbf{H}_2)} \times \frac{\mathbf{P}(\mathbf{H}_1)}{\mathbf{P}(\mathbf{H}_2)}$$

The first factor on the right side of the equation is known as the *likelihood ratio*; the second is the ratio of priors, which measures our relative degree of belief in these two hypotheses before E entered the picture. Since the ratio on the left side of this equation represents the relative degree of belief in these two hypotheses once E is in hand, this equation shows that updating our prior relative degree of belief in these hypotheses (i.e., before the evidence E was factored in) is simply a matter of multiplying the ratio of prior probabilities times the likelihood ratio.

In this way, the likelihood ratio, i.e.,

$$\frac{\mathbf{P}(E \mid \mathbf{H}_1)}{\mathbf{P}(E \mid \mathbf{H}_2)},$$

is said the measure the strength of evidence that E provides for $\mathbf{H}_1$ in relation to $\mathbf{H}_2$. Thus, since we are treating $\mathbf{H}_1$ as a design hypothesis and $\mathbf{H}_2$ as a chance/evolutionary hypothesis, if this ratio is bigger than 1, E favors the design hypothesis and warrants a design inference. On the other hand, if this ratio is less than 1, it favors the chance/evolutionary hypothesis and therefore does not warrant a design inference.

Although at first blush plausible, this Bayesian approach to design inferences is deeply problematic. I discuss its problems at length in an article titled "Design by Elimination vs. Design by Comparison," which is chapter 33 of my book *The Design Revolution*. That article distinguishes my Fisherian approach to design inferences from the Bayesian approach and can be found at http://www.designinference.com/documents/2005.09.Fisher_vs_Bayes.pdf.