Conditional Probability and Expected Value

February 3, 2015

The Probability Axioms

1. **Normality.** The probability of any proposition X is somewhere between 0 and 1.

$$0 \leq \Pr(X) \leq 1 \tag{1}$$

2. **Certainty.** Let Ω be a proposition that is certain to be true.

$$\Pr(\Omega) = 1 \tag{2}$$

3. **Additivity.** If propositions X and Y are *mutually exclusive*, then the probability of their disjunction is equal to the sum of their probabilities.

Two propositions are *mutually exclusive* just in case they cannot *both* be true.

$$\text{If } X\&Y \text{ are mutually exclusive, } \Pr(X \vee Y) = \Pr(X) + \Pr(Y) \tag{3}$$

The Overlap Rule

What is the probability of a disjunction when its disjuncts are *not* mutually exclusive?

Overlap. The probability of a disjunction is equal to the sum of the probabilities of its disjuncts minus the probability its disjuncts' overlap.

$$\Pr(X \vee Y) = \Pr(X) + \Pr(Y) - \Pr(X \wedge Y) \tag{4}$$

We can derive **The Overlap Rule** from the probability axioms (plus the assumption that logically equivalent propositions have the same probability). Here's how [see pg. 60]:

Extra Assumption: If X and Y are *logically equivalent*, then $\Pr(X) = \Pr(Y)$.

1. From *Propositional Logic:* $(X \vee Y)$ is logically equivalent to $((X \wedge Y) \vee (X \wedge \neg Y) \vee (\neg X \wedge Y))$.
2. From *Propositional Logic:* The propositions $(X \wedge Y)$, $(X \wedge \neg Y)$, and $(\neg X \wedge Y)$ are all mutually exclusive.
3. From *Additivity:* $\Pr((X \wedge Y) \vee (X \wedge \neg Y) \vee (\neg X \wedge Y)) = \Pr(X \wedge Y) + \Pr(X \wedge \neg Y) + \Pr(\neg X \wedge Y)$.
4. So, given the assumption that logically equivalent proposition have the same probability, $\Pr(X \vee Y) = \Pr(X \wedge Y) + \Pr(X \wedge \neg Y) + \Pr(\neg X \wedge Y)$.
5. From math,

$$\begin{aligned}\Pr(X \vee Y) &= \Pr(X \wedge Y) + \Pr(X \wedge \neg Y) + \Pr(\neg X \wedge Y) \\ &= \Pr(X \wedge Y) + \Pr(X \wedge \neg Y) + \Pr(\neg X \wedge Y) + \Pr(X \wedge Y) - \Pr(X \wedge Y)\end{aligned}$$

6. From *Propositional Logic* and *Additivity:*

$$\Pr(X) = \Pr(X \wedge Y) + \Pr(X \wedge \neg Y)$$
$$\Pr(Y) = \Pr(X \wedge Y) + \Pr(\neg X \wedge Y)$$

Hence, $\Pr(X \vee Y) = \Pr(X) + \Pr(Y) - \Pr(X \wedge Y)$.

This is the intuitive idea behind **The Overlap Rule.** If the propositions X and Y are not mutually exclusive, then by adding $\Pr(X)$ to $\Pr(Y)$ in order to get $\Pr(X \vee Y)$, we are "double counting" the possibility in which they are *both* true, i.e., $(X \wedge Y)$. To correct for this, we need to subtract out $\Pr(X \wedge Y)$.

Conditional Probability

Let $\Pr(X \mid Y)$ be the probability of X *conditional* on Y.
It is defined as follows:

$\Pr(X|Y)$ is, roughly, the probability that X is the case on the assumption that Y is the case.

$$\Pr(X \mid Y) = \frac{\Pr(X \wedge Y)}{\Pr(Y)} \tag{5}$$

The probability axioms and the definition of conditional probability all hold in conditional form. That is, for a proposition E,

Assume that $\Pr(E) > 0$.

1. **Normality (Conditional Form).**

$$0 \leq \Pr(X \mid E) \leq 1$$

2. **Certainty (Conditional Form).**

$$\Pr(\Omega \mid E) = 1$$

3. **Additivity (Conditional Form).**

   If $X \& Y$ are mutually exclusive, $\Pr(X \vee Y \mid E) = \Pr(X \mid E) + \Pr(Y \mid E)$

4. **Conditional Probability (Conditional Form).** Assume that both $\Pr(E) > 0$ and $\Pr(Y \mid E) > 0$. Then,

$$\Pr(X \mid (Y \wedge E)) = \frac{\Pr((X \wedge Y) \mid E)}{\Pr(Y \mid E)}$$

Proof of 4. Given the definition of conditional probability, we know that

$$\Pr(X \mid (Y \wedge E)) = \frac{\Pr(X \wedge Y \wedge E)}{\Pr(Y \wedge E)}$$

And

$$\Pr(X \wedge Y \wedge E) = \Pr((X \wedge Y)|E) \cdot \Pr(E)$$
$$\Pr(Y \wedge E) = \Pr(Y|E) \cdot \Pr(E)$$

So, we have

$$\frac{\Pr(X \wedge Y \wedge E)}{\Pr(Y \wedge E)} = \frac{\Pr((X \wedge Y)|E) \cdot \Pr(E)}{\Pr(Y|E) \cdot \Pr(E)} = \frac{\Pr((X \wedge Y)|E)}{\Pr(Y|E)}$$

This means that a conditional probability function $\Pr(\bullet \mid E)$ is, itself, a probability function.

More Rules and Definitions

The Multiplication Rule: If $\Pr(E) > 0$, then

$$\Pr(X \wedge E) = \Pr(X \mid E) \cdot \Pr(E) \tag{6}$$

The Total Probability Rule: If $0 < \Pr(E) < 1$, then

$$\Pr(X) = \Pr(X \mid E) \cdot \Pr(E) + \Pr(X \mid \neg E) \cdot \Pr(\neg E) \tag{7}$$

The Logical Consequence Rule: Suppose that Y *logically entails* X. Then

$$\Pr(Y) \leq \Pr(X) \tag{8}$$

Statistical Independence. X and Y are said to be *statistically independent* just in case $\Pr(X \mid Y) = \Pr(X)$.

If X and Y are statistically independent, $\Pr(X \wedge Y) = \Pr(X) \cdot \Pr(Y)$.

Proof. From **The Multiplication Rule:**

$$\Pr(X \wedge Y) = \Pr(X \mid Y) \cdot \Pr(Y)$$

And, from the definition of **Statistical Independence:**

$$\Pr(X \mid Y) = \Pr(X)$$

So, $\Pr(X \wedge Y) = \Pr(X) \cdot \Pr(Y)$.

Bayes' Rule

Let H be some hypothesis. And let E be some evidence.

Bayes' Rule. Assume that $\Pr(E) > 0$. Then,

$$\Pr(H \mid E) = \frac{\Pr(E \mid H) \cdot \Pr(H)}{\Pr(E)} \tag{9}$$

$$= \frac{\Pr(E \mid H) \cdot \Pr(H)}{\Pr(E \mid H) \cdot \Pr(H) + \Pr(E \mid \neg H) \cdot \Pr(\neg H)} \tag{10}$$

This rule follows from the definition of conditional probability.

Proof. From the definition of conditional probability, we have that

$$\Pr(E \mid H) = \frac{\Pr(E \wedge H)}{\Pr(H)}$$

$$\Pr(H \mid E) = \frac{\Pr(E \wedge H)}{\Pr(E)}$$

So, $\Pr(E \mid H) \cdot \Pr(H) = \Pr(E \wedge H)$. And thus

$$\Pr(H \mid E) = \frac{\Pr(E \mid H) \cdot \Pr(H)}{\Pr(E)}$$

Example Problem 1: Spiders.

Let G be the proposition that *the bananas are from Guatemala.* Let H be the propositions that *the bananas are from Honduras.* And let T be the propositions that *the bananas had a tarantula on them.* Given that we've found a tarantula in the bananas, what's the probability that they came from Guatemala?

$$\Pr(G \mid T) = \frac{\Pr(T \mid G) \cdot \Pr(G)}{\Pr(T \mid G) \cdot \Pr(G) + \Pr(T \mid H) \cdot \Pr(H)}$$

$$= \frac{.06 \times .6}{(.06 \times .6) + (.03 \times .4)}$$

$$= \frac{.036}{.036 + .012} = \frac{.036}{.048} = .75$$

$$\Pr(T \mid H) = .03$$
$$\Pr(T \mid G) = .06$$
$$\Pr(H) = .4$$
$$\Pr(G) = .6$$

Example Problem 2: base rate fallacy

Let B be the proposition that *it was a blue cab.* Let R be the proposition that *it was a red cab.* Let "B" be the proposition that *the witness said it was a blue cab.* And let "R" be the proposition that *the witness said it was a red cab.* Given that the witness said it was a blue cab, what's the probability that it was a blue cab?

$$\Pr(B \mid "B") = \frac{\Pr("B" \mid B) \cdot \Pr(B)}{\Pr("B" \mid B) \cdot \Pr(B) + \Pr("B" \mid R) \cdot \Pr(R)}$$

$$= \frac{.9 \times .01}{(.9 \times .01) + (.1 \times .99)} = \frac{.009}{.009 + .099} = \frac{9}{108} \approx .083$$

$$\Pr("X" \mid X) = .9$$
$$\Pr(R) = .99$$
$$\Pr(B) = .01$$

Expected Value

What is the value of performing an act when you are uncertain about what would happen were you to perform it?

1. You are confronted with a range of different possible *acts*, $\mathbf{A_1}, \mathbf{A_2}, \ldots, \mathbf{A_n}$, which are mutually exclusive and exhaustive.
2. For each possible act, consider a (finite) set of mutually exclusive and exhaustive "possible consequences": $C_1, C_2, \ldots, C_k$.
3. For each possible consequence C_i and act $\mathbf{A}$, we assign to it a *utility:* $U(C_i \wedge \mathbf{A})$
4. The *expected value* of an act $\mathbf{A}$ is found by multiplying the utility by the conditional probability for each consequence, and then adding them all up.

$$Exp(\mathbf{A}) = \sum_{i=1}^{n} \Pr(C_i \mid \mathbf{A}) \cdot U(C_i \wedge \mathbf{A}) \tag{11}$$

The expected utility of an act is a *weighted average:* it's the average utility of a possible consequence, weighted by the probability of that consequence coming about.

When a set of propositions are mutually exclusive (i.e., at most one of them is true) and exhaustive (i.e., at least one of them is true), we say that the set forms a *partition.*

The average of $a_1, \ldots, a_n$ is

$$\frac{a_i + \ldots a_n}{n} = \sum_{i=1}^{n} (\frac{1}{n}) \cdot a_i$$

Here, the "weights" are all the same. We can get a weighted average by changing the weights (just so long as they sum to 1).

Potential Decision Rules

How should we choose which act, from the set of all available ones, to perform?

Proposal: Maximize Expected ($) Value.

OBJECTION: We value other things besides money.

Proposal: Postulate the existence of *utiles* ("unites of pure utility"). Maximize Expected Utiles.

Note: the same dollar amount might be worth different amounts of utiles for different people, in different situations. In fact, it seems like money has *diminishing marginal utility.*

OBJECTION 2: What about Risk Aversion?

Example. I'm going to toss a fair coin. And I offer you the following two deals.

$\mathbf{A_1}$: No matter how the coin lands, you get $50.

$\mathbf{A_2}$: If the coin lands *Heads*, you get $0; if the coin lands *Tails*, you get $100.

Notice that $Exp(\mathbf{A_1}) = Exp(\mathbf{A_2})$.

Suppose you give every dollar equal value. Is it irrational for you to prefer $\mathbf{A_1}$ to $\mathbf{A_2}$?

Is this just a problem for the first proposal? Or is this a problem for both proposals?

Suggestion: Allow the utility function to take account of things like *risk* and *uncertainty.*

OBJECTION: Ad hoc? It collapses things that should be kept seperate?