R T Cox: Probability, Frequency and Reasonable Expectation
AJP 14 1 (1946)

To choose the idea of frequency in an ensemble or the idea of reasonable expectation as the primary meaning of probability has distinguished the two main schools of thought in the theory.

Probability is recognized also as providing a measure of the reasonable expectation of an event in a single trial. According to the second main school of probability, this measure of reasonable expectation, rather than the frequency in an ensemble, is the primary meaning of probability.

There are probabilities in the sense of reasonable expectations for which no ensemble exists and for which, if one is conceived, it is clearly no more than a convenient mental artifice.

**Relations of Reasonable Expectation Consistent with Symbolic Logic**

Letters, a, b, c, · · ·, will denote propositions. We assume the first order logic with ¬, ∧ and ∨:

1. ¬¬a = a
2. a ∨ b = b ∨ a [a ∧ b = b ∧ a]
3. a ∨ a = a [a ∧ a = a]
4. a ∨ (b ∨ c) = (a ∨ b) ∨ c = a ∨ b ∨ c [∧ version as well]
5. (de Morgan) ¬(a ∨ b) = ¬a ∧ ¬b [¬(a ∧ b) = ¬a ∨ ¬b]
6. a ∧ (a ∨ b) = a [a ∨ (a ∧ b) = a]

Definition: b|a: a measure of reasonable credibility (mrc) of b when a is known.

If b|a is a mrc, any function of it is also an mrc. a and b may have no relation.

Axiom 1: c ∧ b|a = F(c|b ∧ a, b|a), where F is some function of two variables.

We use (4) to make a functional equation for F. Notice that there are two ways to analyze d ∧ c ∧ b|a = (d ∧ c) ∧ b|a = d ∧ (c ∧ b)|a:

\[
\begin{align*}
    d ∧ c ∧ b|a &= F(d ∧ c ∧ b ∨ a, b|a) = F(F(d|c ∧ b ∧ a, c|b ∧ a), b|a) \\
    &= F(d|c ∧ b ∧ a, c ∧ b|a) = F(d|c ∧ b ∧ a, F(c|b ∧ a, b|a)).
\end{align*}
\] (0.0.1)

Notes:

1. Richard Threlkeld Cox (August 5, 1898 - May 2, 1991) was a professor of physics at Johns Hopkins University, known for Cox’s theorem, his most important work, relating to the foundations of probability. Cox’s theorem is a derivation of the laws of probability theory from a certain set of postulates. This derivation justifies the so-called "logical" interpretation of probability. As the laws of probability derived by Cox’s theorem are applicable to any proposition, logical probability is a type of Bayesian probability. Other forms of Bayesianism, such as the subjective interpretation, are given other justifications. [Wikipedia: Cox & Cox’s theorem]

2. The significant point of the concept of ensemble is that the initial circumstances are assumed to be capable of indefinite repetition, these repetitions constituting the ensemble.

3. For example, when the probability is calculated that more than one planetary system exists in the universe, it is barely tenable even as an artifice that this refers to the number of universes having more than one planetary system among an indefinitely large number of universes, all resembling in some way the universe, which by definition is all-inclusive.

4. For example, a= “Caesar invaded Britain” and b = “Tomorrow will be warmer than today”.

Thus,
\[ F(F(x, y), z) = F(x, F(y, z)). \] (0.0.3)

If \( F \) is twice differentiable, we may show that there are a constant \( C \) and a function \( f \) such that
\[ C f(F(x, y)) = f(x)f(y). \] (0.0.4)

Therefore, \( c \wedge b|a = F(c|b \wedge a, b|a) \) implies
\[ C f(c \wedge b|a) = f(c|b \wedge a)f(b|a). \] (0.0.5)

Since \( b|a \) and \( f(b|a) \) are mrc, the choice of \( f \) is purely a matter of convention. If we set \( c \) to be \( b \), \( c \wedge b|a = b|a \), so (0.0.5) implies
\[ C = f(b|b \wedge a). \] (0.0.6)

\( b \wedge a \) implies \( b \), so \( C \) is the mrc of certainty. Let us set \( C = 1 \). Thus,
\[ c \wedge b|a = c|b \wedge a b|a \] (0.0.7)

cf \( P(C \cap B \mid A) = P(C \mid B \cap A)P(B \mid A) \). However, \( (c \wedge b|a)^m = (c|b \wedge a)^m (b|a)^m \) is also true for any \( m \), so we cannot yet uniquely relate mrc to \( P \).

**Axiom 2:** \( \neg b|a = S(b|a) \) for some function \( S \).

This is reasonable, since \( \neg b \) is determined when \( b \) is specified. The \( S \) must be an involution since \( \neg \neg a = a \):
\[ S(S(x)) = x. \] (0.0.8)

We wish to get rid of \( \neg \): \( \neg(c \vee b)|a = (\neg c \wedge \neg b)|a = \neg c \wedge \neg b \wedge \neg b|a, \) so
\[ S(c \vee b|a) = S(c|\neg b \wedge a)S(b|a). \] (0.0.9)

Or,
\[ S(c|\neg b \wedge a) = S(c \vee b|a)S(b|a). \] (0.0.10)

Since \( S \) is involutive
\[ c|\neg b \wedge a = S(S(c \vee b|a)/S(b|a)). \] (0.0.11)

Notice that \( \neg b \wedge c|a = c|\neg b \wedge a \neg b|a \), so
\[ c|\neg b \wedge a = \neg b \wedge c|a/\neg b|a = \neg b|c \wedge a c|a/\neg b|a = S(b|c \wedge a)c|a/S(b|a) \] (0.0.12)

Therefore, from (0.0.11) and (0.0.12)
\[ S(b|c \wedge a)c|a/S(b|a) = S(S(c \vee b|a)/S(b|a)). \] (0.0.13)

Since \( b|c \wedge a = c \wedge b|a/c|a \), this equation now reads
\[ S(c \wedge b|a/c|a)c|a/S(b|a) = S(S(c \vee b|a)/S(b|a)). \] (0.0.14)

or
\[ S(c \wedge b|a/c|a)c|a = S(S(c \vee b|a)/S(b|a))S(b|a). \] (0.0.15)
Now, replacing $b$ with $c \land d$, we have (notice $c \land (c \land d) = c \land d$ and $c \lor (c \lor d) = c$)

$$S(c \land d|a/c|a)c|a = S(S(c|a)/S(c \land d|a))S(c \land d|a). \tag{0.0.16}$$

This has a highly symmetric form as can be recognized by setting $c|a = x$ and $S(c \land d|a) = y$ (note also $c \land d|a = S(S(c \land d|a)) = S(y)$)

$$S(y)/x \times S(x)/y = \tag{0.0.17}$$

If $S$ is twice differentiable, then

$$S(x) = (1 - x^m)^{1/m}. \tag{0.0.18}$$

Axiom 2 implies

$$(b|a)^m = S(-b|a)^m = 1 - (-b|a)^m \tag{0.0.19}$$

Since $m$ can be anything, let choose $m = 1$:

$$b|a - b|a = 1. \tag{0.0.20}$$

If we choose $b = a$, this reads

$$a|a + a|a = 1 \tag{0.0.21}$$

Here, two mrc are certainty and impossibility. Since certainty is given mrc = 1, impossibility should have mrc = 0.

We can demonstrate

$$c \land b|a + \neg c \land b|a = b|a. \tag{0.0.22}$$

and

$$c \lor b|a = c|a + b|a - c \land b|a. \tag{0.0.23}$$

About the functional equations:

For $F(x, F(y, z)) = F(F(x, y), z)$, obviously $F(x, y) = Cxy$ satisfies the equation.

For $S(S(y)/x)x = S(S(x)/y)y$, obviously $S(x) = 1 - x$ satisfies the equation. If $S$ is involutive and if it satisfies this equation, then this must be true for $x$ replaced by any function of $x$. There is such a function. The question is whether any solution can be reduced to $1 - x$. 
