# Significance Tests, Belief Calculi, and Burden of Proof in Legal and Scientific Discourse

Julio Michael Stern
BioInfo and Computer Science Department
University of São Paulo, Brazil.

The FBST Value of Evidence
Full Bayesian Significance Test
(Pereira and Stern, 1999)

Posterior desnsity, likelihood and prior:

$$p_x(\theta) \propto L(\theta \,|\, x)\, p(\theta).$$

Null hypothesis:

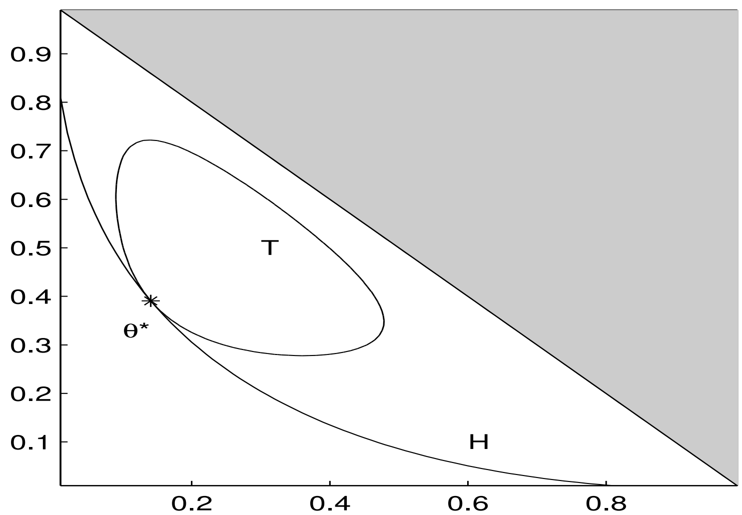
$$\Theta_H = \{\theta \in \Theta \,|\, g(\theta) \leq \mathbf{0} \wedge h(\theta) = \mathbf{0}\}$$

Sharp (precise) hypotheses:
$\dim(\Theta_H) < \dim(\Theta)$.

Evidence against the hypothesis:

$$\begin{aligned}
\text{Ev}(H) &= \int_{T_H} p_x(\theta)\, d\theta \ , \text{ where} \\
T_H &= \{\theta \in \Theta \mid s(\theta) > s_H\} \\
s_H &= \sup_{\theta \,\in\, \Theta_H} s(\theta) \\
s(\theta) &= \left(\frac{p_x(\theta)}{r(\theta)}\right)
\end{aligned}$$

$s(\theta)$ is the Posterior Surprise
If the reference density $r(\theta) \propto 1$ ,
the Tangent set $T_H$ , HRSS = HDPS




Hardy-Weinberg genetic equilibrium:
$n$ sample size, $x_1, x_3$ homozygote counts
$x_2 = n - x_1 - x_3$ heterozygote count

$$
\begin{aligned}
p_x(\theta \mid x) &\propto \theta_1^{x_1}\theta_2^{x_2}\theta_3^{x_3}\ , \quad r(\theta) \propto 1 \\
\Theta &= \{\theta \geq 0 \mid \theta_1 + \theta_2 + \theta_3 = 1\} \\
\Theta_H &= \{\theta \in \Theta \mid \theta_3 = (1 - \sqrt{\theta_1})^2\}
\end{aligned}
$$

ABC - Abstract Belief Calculus, Darwiche and Ginsberg, generalization of Probability calc.

Abstract Support Function, $\Phi$
coherence conditions on support states:

**A1:** Equivalent statements must have the same support,

$$(A \Leftrightarrow B) \Rightarrow \Phi(A) = \Phi(B)$$

**A2:** Support Summation,

$$\oplus : \Phi(\mathcal{U}) \times \Phi(\mathcal{U}) \mapsto \Phi(\mathcal{U})$$

support of disjunction of two logically disjoint statements is the sum of their individual support values,

$$\neg(A \wedge B) \Rightarrow \Phi(A \vee B) = \Phi(A) \oplus \Phi(B)$$

**A3:** If A implies B, which implies C, and A and C have the same support, then all three statements have the same support,

$$((A \Rightarrow B \Rightarrow C) \wedge (\Phi(A) = \Phi(C)))$$

$$\Rightarrow \Phi(B) = \Phi(A)$$

**A4:** False statements have zero support value,

$$A \text{ false } \Rightarrow \Phi(A) = 0$$

**A5:** Tautological statements have full support,

$$A \text{ true } \Rightarrow \Phi(A) = 1$$

Axioms A1 to A5 $\Rightarrow$ algebraic properties:

**X0:** Symmetry,

$$a \oplus b = b \oplus a$$

**X1:** Transitivity,

$$(a \oplus b) \oplus c = a \oplus (b \oplus c)$$

**X2:** Convexity,

$$\text{if } a \oplus b \oplus c = a \quad \text{then } a \oplus b = a$$

**X3:** There is a unique element 0

$$\forall a \in \Phi(\mathcal{U}), a \oplus 0 = a$$

**X4:** There is a unique element $1 \neq 0$

$$\forall a \in \Phi(\mathcal{U}), \exists!\, b \in \Phi(\mathcal{U}) \mid a \oplus b = 1$$

Support function and summation, $\langle \Phi, \oplus \rangle$ is a Partial Support Structure. Examples:

| $\Phi(\mathcal{U})$ | $a \oplus b$ | 0 | 1 | $a \preceq b$ | Calculus |
|---|---|---|---|---|---|
| $\{0, 1\}$ | $\max(a, b)$ | 0 | 1 | $a \leq b$ | Cls. Logic |
| $[0, 1]$ | $a + b$ | 0 | 1 | $a \leq b$ | Probablty |
| $[0, 1]$ | $\max(a, b)$ | 0 | 1 | $a \leq b$ | Possiblty |
| $\{0..\infty\}$ | $\min(a, b)$ | $\infty$ | 0 | $b \leq a$ | Disbelief |

The support value of a statement does not determine the support value of its negation, but the belief value of a statement, $\ddot{\Phi}$, does,

$$\ddot{\Phi}(A) = \langle \Phi(A), \Phi(\neg A) \rangle$$

Partial support structures also define partial orders on $\Phi(\mathcal{U})$, $\preceq$, and on $\ddot{\Phi}(\mathcal{U})$, $\sqsubseteq$.

$$a \preceq b \Leftrightarrow \exists\, c \,|\, a \oplus c = b$$

$$\langle a, b \rangle \sqsubseteq \langle c, d \rangle \Leftrightarrow a \preceq c \text{ and } d \preceq b$$

The extreme, minimal and maximal, states of support and belief are, 0 and 1 for the support order, and $\langle 0, 1 \rangle$ and $\langle 1, 0 \rangle$ for the belief order. Statements with minimal and maximal belief are Rejected and Accepted.

Evidence and Onus Probandi
Support value, $\Phi(H) = \overline{\mathsf{Ev}}(H) = 1 - \mathsf{Ev}(H)$, for a hypothesis, $H : \theta \in \Theta_H \subseteq \Theta$.

**R1, Value of Evidence as a Probability:**

$$\mathsf{Ev}(H) = \int_{\Gamma_H} p_x(\theta) d\theta$$

If $\theta \in \Gamma_H$, $\theta$ "constitutes evidence against $H$"

If $\theta \in \Theta_H$, $\theta$ "is compatible with (admissible, legal or valid by) $H$ "

**R2, Relative Surprise:** Whether a parameter point $\theta$ constitutes or not evidence against $H$ depends only on the order in the parameter space established by the value of the posterior surprise relative to a given reference density,

$$s(\theta) = p_x(\theta)/r(\theta)$$

**R3, No Invalid (Self) Incrimination:** If $\theta$ is compatible with $H$, it can not constitute evidence against $H$,

$$\Theta_H \cap \Gamma_H = \emptyset$$

**R4, De Morgan's Law:** A point $\theta$ constitutes evidence against $H$ iff it constitutes evidence against all of its terms,

$$\text{if } H = A \vee B \text{ then } \Gamma_H = \Gamma_A \cap \Gamma_B$$

**R5, Most Favorable Interpretation:** The evidence in favor of a composite hypothesis is the most favorable evidence in favor of its terms,

$$\text{if } H = A \vee B \text{ then}$$

$$\overline{\text{Ev}}(H) = \max(\overline{\text{Ev}}(A), \overline{\text{Ev}}(B))$$

**R6, Coherent Support:** $\langle \overline{\text{Ev}}, \text{max} \rangle$ must be a partial support structure.

**R7, Continuity:** If $p_x(\theta)$, $g(\theta)$ and $h(\theta)$ are smooth (continuous, differentiable, etc.) functions on its arguments, then so is $\text{Ev}(H)$.

**R8, Invariance:** $\text{Ev}(H)$ is invariant under bijective smooth reparameterizations, of the:
a) parameter space.
b) hypothesis representation.

**R9, Consistency:** As sample size $\to \infty$, $\text{Ev}(H)$ converges to 0 or 1, according to whether H is true or false

Onus Probandi Principle or Burden of Proof, or Safe Harbor Liability Rule:

"There is no liability as long as there is a reasonable basis for belief, effectively placing the burden of proof (Onus Probandi) on the plaintiff, who, in a lawsuit, must prove false a defendant's misstatement, without making any assumption not explicitly stated by the defendant, or tacitly implied by an existing law or regulatory requirement."

"Moreover, the party against whom the motion is directed is entitled to have the trial court construe the evidence in support of its claim as truthful, giving it its most favorable interpretation, as well as having the benefit of all reasonable inferences drawn from that evidence."

Conditionalization, given by a Support Scaling:

$$\oslash : \Phi(\mathcal{U}) \times \Phi(\mathcal{U}) \mapsto \Phi(\mathcal{U})$$

$\Phi(B)$ and $\Phi_A(B)$, are the unconditional and conditional support value of B given A.

**A6:** The conditional support value of $B$ given $A \vee B$ is a function of the unconditional support values of $B$ and $A \vee B$,

$$\Phi_{A\vee B}(B) = \Phi(B) \oslash \Phi(A \vee B)$$

It can be seen that axiom 6 is equivalent to

$$\Phi_A(B) = \Phi(A \wedge B) \oslash \Phi(A)$$

**A7:** Accepting a non-rejected statement retains all accepted statements,

$$(\Phi(A) \neq 0 \;\wedge\; \Phi(B) = 0) \;\Rightarrow\; \Phi_A(B) = 0$$

**A8:** Accepting an accepted statement does not change the conditional support function,

$$\Phi(A) = 1 \;\Rightarrow\; \Phi_A = \Phi$$

**A9:** When $A \vee B$ is equally supported by two support functions, conditioning on $A \vee B$ does not introduce equality or order between the unconditional supports of $A$,

$$\text{if } \Phi(A \vee B) = \Psi(A \vee B) \text{ then}$$

$$\Phi_{A\vee B}(A) \preceq (=) \Psi_{A\vee B}(A) \Rightarrow \Phi(A) \preceq (=) \Psi(A)$$

**A10:** After accepting the logical consequences of a statement, $A$, the conditional support of $A$ either increases or does not change,

$$\Phi(A \vee B) \neq 0 \ \Rightarrow \ \Phi(A) \preceq \Phi_{A\vee B}(A)$$

**A11:** If the conditional support of $A$ given $C$ equals its conditional support given $B \wedge C$, then the conditional support of $B$ given $C$ equals its conditional support given $A \wedge C$,

$$(\Phi(A \wedge B \wedge C) \neq 0 \wedge \Phi_C(A) = \Phi_{B\wedge C}(A))$$
$$\Rightarrow \ \Phi_C(B) = \Phi_{A\wedge C}(B)$$

$\langle \Phi(\mathcal{U}), \oplus, \oslash \rangle$ is a Support Structure.

For the former examples, the scaling functions are:

$$\Phi_A(B) = \min\left(\Phi(A \wedge B), \Phi(A)\right)$$

for classical logic;

$$\Phi_A(B) = \frac{\Phi(A \wedge B)}{\Phi(A)}$$

for probability and possibility calculus; and

$$\Phi_A(B) = \Phi(A \wedge B) - \Phi(A)$$

for disbelief calculus.

## Coexistent Belief Calculi

A critical interpretation of FBST's value of evidence, in the context set by the previous sections, can help us elucidate the benefits and some apparent paradoxes of using the FBST in statistical testing.

The FBST support values $\overline{\text{Ev}}(H)$, are computed using standard probability calculus on $\Theta$ which has an intrinsic conditionalization operator. The computed evidences form a possibilistic partial support structure. Therefore, two belief calculi are in simultaneous use in the Full Bayesian Testing setup: probability and possibility calculus.

Darwiche and Ginsberg make some interesting remarks concerning support and belief orders. Namely:

1- If two statements are equally believed, then they are equally supported; but not the converse.

2- Rejected statements are always minimally supported, and accepted statements are always maximally supported. But although minimally supported statements are rejected, maximally supported sentences are not necessarily accepted.

3- A statement and its negation may be maximally supported at the same time, while neither of them may be accepted.

Consider the hypotheses

$$A : \theta \in \Theta \text{ and } B : \theta \in \{\widehat{\theta}\}$$

where $\widehat{\theta}$ is the unique maximizer of a smooth proper posterior density in the parameter space

$$\Theta = \mathcal{R}^p \ , \ \{\widehat{\theta}\} = \arg\max_{\theta \in \Theta} p_x(\theta).$$

Asume a uniform reference, $r(\theta) \propto 1$. We have,

$$\overline{\text{Ev}}(A) = \overline{\text{Ev}}(B) = \overline{\text{Ev}}(\neg B) = 1$$

$$\text{and} \quad \overline{\text{Ev}}(\neg A) = 0$$

So both A and B have full support, but A is accepted, while B is not.

This example, or variations of it, were given to the author as either an example of how a support function should work in the juridical context, or as a FBST paradox, in the context of traditional statistical tests of significance.

In the juridical context, the interpretation is as follows: A defendant describes a system (machine, software, genetic code etc.) by a parameter $\theta$, and claims that $\theta$ has been set to a value in a legal or valid null set, $\Theta_H$. The parameter can not be observed directly, but we can observe a random variable whose distribution is a function $f(x;\theta)$. The parameter $\theta$ has been set to one, and only one value. Claiming that $\theta$ has been set at the most likely value, $\theta = \widehat{\theta}$, (given $n$ observed outcomes) must give the defendant's claim full support, for being absolutely vague, i.e., claiming only that $\theta \in \Theta$, cannot put him in a better position.

In most traditional statistical tests of significance, $\Phi(\Theta_H)$ is a probability measure of the null set, $\Pr(\Theta_H)$. If $\Theta_H$ is a singleton in $\mathcal{R}^p$, with a smooth posterior, then it should have null support. Indeed, the refutation of any sharp hypothesis is a price many philosophers, see (Popper 1989), and most statisticians are ready to pay, as explicitly stated by I.J.Good:

"If by the truth of Newtonian mechanics we mean that it is approximately true in some appropriate well defined sense we could obtain strong evidence that it is true; but if we mean by its truth that it is exactly true then it has already been refuted. ... Very often the statistician doesn't bother to make it quite clear whether his null hypothesis is intended to be sharp or only approximately sharp. ... It is hardly surprising then that many Fisherians (and Popperians) say that - you can't get (much) evidence in favor of the null hypothesis but can only refute it."