A Better Bayesian Convergence Theorem

Notation

hypotheses: $h_i \quad h_j \quad \ldots$

background and auxiliaries: b

experimental/observation conditions:

$$c_1 , c_2 , \ldots , c_n \ : \ c^n$$

evidential outcomes:

$$e_1 , e_2 , \ldots , e_n \ : \ e^n$$

likelihoods: $P[e^n \mid h_i{\cdot}b{\cdot}c^n]$

Notation

Likelihoods Priors Posteriors

$P[e \mid h_i] \quad : \quad P_\alpha[h_i] \quad \Rightarrow \quad P_\alpha[h_i \mid e]$

$P[e \mid h_i \cdot b \cdot c]: \quad P_\alpha[h_i \mid b] \Rightarrow P_\alpha[h_i \mid b \cdot c \cdot e]$

$$P_\alpha[h_i \mid b \cdot c^n \cdot e^n]$$

$$= \frac{P[e^n \mid h_i \cdot b \cdot c^n] \cdot P_\alpha[h_i \mid b]}{P_\alpha[e^n \mid b \cdot c^n]} \cdot \frac{P_\alpha[c^n \mid h_i \cdot b]}{P_\alpha[c^n \mid b]}$$

$$= \frac{P[e^n \mid h_i \cdot b \cdot c^n] \cdot P_\alpha[h_i \mid b]}{P_\alpha[e^n \mid b \cdot c^n]}$$

$$\frac{P_\alpha[h_j \mid b\cdot c^n\cdot e^n]}{P_\alpha[h_i \mid b\cdot c^n\cdot e^n]}$$

$$= \frac{P[e^n \mid h_j\cdot b\cdot c^n]}{P[e^n \mid h_i\cdot b\cdot c^n]} \cdot \frac{P_\alpha[h_j \mid b]}{P_\alpha[h_i \mid b]} \cdot \frac{P_\alpha[c^n \mid h_j\cdot b]}{P_\alpha[c^n \mid h_i\cdot b]}$$

$$= \frac{P[e^n \mid h_j\cdot b\cdot c^n]}{P[e^n \mid h_i\cdot b\cdot c^n]} \cdot \frac{P_\alpha[h_j \mid b]}{P_\alpha[h_i \mid b]}$$

$$\Omega_\alpha[\sim h_i \mid b\cdot c^n\cdot e^n] \;=\; \frac{P_\alpha[\sim h_i \mid b\cdot c^n\cdot e^n]}{P_\alpha[h_i \mid b\cdot c^n\cdot e^n]}$$

$$= \sum_{j\neq i} \frac{P[e^n \mid h_j\cdot b\cdot c^n]}{P[e^n \mid h_i\cdot b\cdot c^n]} \cdot \frac{P_\alpha[h_j \mid b]}{P_\alpha[h_i \mid b]}$$

$$+ \frac{P_\alpha[e^n \mid h_K\cdot b\cdot c^n]}{P[e^n \mid h_i\cdot b\cdot c^n]} \cdot \frac{P_\alpha[h_K \mid b]}{P_\alpha[h_i \mid b]}$$

where h_K is the catch-all, “something-else” hypothesis.

$$P_\alpha[h_i \mid b\cdot c^n\cdot e^n] = \frac{1}{1 + \Omega_\alpha[{\sim}h_i \mid b\cdot c^n\cdot e^n]}$$

$$\Sigma_{j\neq i} \quad \frac{P[e^n \mid h_j \cdot b \cdot c^n]}{P[e^n \mid h_i \cdot b \cdot c^n]} \cdot \frac{P_\alpha[h_j \mid b]}{P_\alpha[h_i \mid b]}$$

$$\leq \qquad \Omega_\alpha[\sim h_i \mid b \cdot c^n \cdot e^n] \qquad \leq$$

$$\Sigma_{j\neq i} \quad \frac{P[e^n \mid h_j \cdot b \cdot c^n]}{P[e^n \mid h_i \cdot b \cdot c^n]} \cdot \frac{P_\alpha[h_j \mid b]}{P_\alpha[h_i \mid b]}$$

$$+ \quad \frac{1}{P[e^n \mid h_i \cdot b \cdot c^n]} \cdot \frac{P_\alpha[h_K \mid b]}{P_\alpha[h_i \mid b]}$$

$$\frac{P_\alpha[h_j \mid b\cdot c^n\cdot e^n]}{P_\alpha[h_i \mid b\cdot c^n\cdot e^n]} = \frac{P[e^n \mid h_j\cdot b\cdot c^n]}{P[e^n \mid h_i\cdot b\cdot c^n]} \cdot \frac{P_\alpha[h_j \mid b]}{P_\alpha[h_i \mid b]}$$

Sufficient Conditions for the ***Likely*** Bayesian Refutation

of False Alternatives to the True Hypothesis

choose any $\varepsilon > 0$
consider the set of outcome streams:

$$\{e^n : P[e^n \mid h_j \cdot b \cdot c^n] / P[e^n \mid h_i \cdot b \cdot c^n] < \varepsilon\}$$

----

now consider the sentence:

$$\vee\{e^n : P[e^n \mid h_j \cdot b \cdot c^n] / P[e^n \mid h_i \cdot b \cdot c^n] < \varepsilon\}$$

Consider

$$P[\vee\{e^n : P[e^n \mid h_j \cdot b \cdot c^n]/P[e^n \mid h_i \cdot b \cdot c^n] < \varepsilon\} \mid h_i \cdot b \cdot c^n]$$

The Outcome Space

c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9
$o_{1,1}$	$o_{2,1}$	$o_{3,1}$	$o_{4,1}$	$o_{5,1}$	$o_{6,1}$	$o_{7,1}$	$o_{8,1}$	…
$o_{1,2}$	$o_{2,2}$	$o_{3,2}$	$o_{4,2}$	$o_{5,2}$	$o_{6,2}$	$o_{7,2}$	$o_{8,2}$	…
$o_{1,3}$	$o_{2,3}$	$o_{3,3}$	$o_{4,3}$	$o_{5,3}$	$o_{6,3}$	$o_{7,3}$	$o_{8,3}$	…
$o_{1,4}$	$o_{2,4}$	$o_{3,4}$	$o_{4,4}$	$o_{5,4}$	$o_{6,4}$	$o_{7,4}$	$o_{8,4}$	…
$o_{1,5}$	$o_{2,5}$	$o_{3,5}$	$o_{4,5}$	$o_{5,5}$		$o_{7,5}$	$o_{8,5}$	…
$o_{1,6}$	$o_{2,6}$	$o_{3,6}$	$o_{4,6}$	$o_{5,6}$		$o_{7,6}$	$o_{8,6}$	…
$o_{1,7}$		$o_{3,7}$	$o_{4,7}$	$o_{5,7}$		$o_{7,7}$	$o_{8,7}$	…
$o_{1,8}$			$o_{4,8}$	$o_{5,8}$		$o_{7,8}$	$o_{8,8}$	…
…			…	…	…	…	…	…

for each h,

$$P[o_{ku} \cdot o_{kv} \mid h \cdot b \cdot c_{k}] = 0$$

$$\sum_{u=1}^{w} P[o_{ku} \mid h \cdot b \cdot c_k] = 1$$

Possible Path of Evidence Stream through the Outcome Space

c^n

	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9
	$o_{1,1}$	$o_{2,1}$	$o_{3,1}$	$o_{4,1}$	$o_{5,1}$	$o_{6,1}$	$o_{7,1}$	$o_{8,1}$	…
	$o_{1,2}$	$o_{2,2}$	$o_{3,2}$	$o_{4,2}$	$o_{5,2}$	$o_{6,2}$	$o_{7,2}$	$o_{8,2}$	…
	$o_{1,3}$	$o_{2,3}$	$o_{3,3}$	$o_{4,3}$	$o_{5,3}$	$o_{6,3}$	$o_{7,3}$	$o_{8,3}$	…
	$o_{1,4}$	$o_{2,4}$	$o_{3,4}$	$o_{4,4}$	$o_{5,4}$	$o_{6,4}$	$o_{7,4}$	$o_{8,4}$	…
e^n	$o_{1,5}$	$o_{2,5}$	$o_{3,5}$	$o_{4,5}$	$o_{5,5}$		$o_{7,5}$	$o_{8,5}$	…
	$o_{1,6}$	$o_{2,6}$	$o_{3,6}$	$o_{4,6}$	$o_{5,6}$		$o_{7,6}$	$o_{8,6}$	…
	$o_{1,7}$		$o_{3,7}$	$o_{4,7}$	$o_{5,7}$		$o_{7,7}$	$o_{8,7}$	…
	$o_{1,8}$			$o_{4,8}$	$o_{5,8}$		$o_{7,8}$	$o_{8,8}$	…
	…			…	…	…	…	…	…

Possible Path of Evidence Stream through the Outcome Space

e^n

$h_i \cdot b$

c^n

c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9
$o_{1,1}$	$o_{2,1}$	$o_{3,1}$	$o_{4,1}$	$o_{5,1}$	$o_{6,1}$	$o_{7,1}$	$o_{8,1}$	…
$o_{1,2}$	$o_{2,2}$	$o_{3,2}$	$o_{4,2}$	$o_{5,2}$	$o_{6,2}$	$o_{7,2}$	$o_{8,2}$	…
$o_{1,3}$	$o_{2,3}$	$o_{3,3}$	$o_{4,3}$	$o_{5,3}$	$o_{6,3}$	$o_{7,3}$	$o_{8,3}$	…
$o_{1,4}$	$o_{2,4}$	$o_{3,4}$	$o_{4,4}$	$o_{5,4}$	$o_{6,4}$	$o_{7,4}$	$o_{8,4}$	…
$o_{1,5}$	$o_{2,5}$	$o_{3,5}$	$o_{4,5}$	$o_{5,5}$		$o_{7,5}$	$o_{8,5}$	…
$o_{1,6}$	$o_{2,6}$	$o_{3,6}$	$o_{4,6}$	$o_{5,6}$		$o_{7,6}$	$o_{8,6}$	…
$o_{1,7}$		$o_{3,7}$	$o_{4,7}$	$o_{5,7}$		$o_{7,7}$	$o_{8,7}$	…
$o_{1,8}$			$o_{4,8}$	$o_{5,8}$		$o_{7,8}$	$o_{8,8}$	…
…			…	…	…	…	…	…

$$P[\vee\{e^n : P[e^n \mid h_j \cdot b \cdot c^n]/P[e^n \mid h_i \cdot b \cdot c^n] < \varepsilon\} \mid h_i \cdot b \cdot c^n]$$

Independent Evidence Assumptions:

1. $P[e^k \mid h_j \cdot b \cdot c_{k+1} \cdot c^k] = P[e^k \mid h_j \cdot b \cdot c^k]$;

2. $P[e_{k+1} \mid h_j \cdot b \cdot c_{k+1} \cdot c^k \cdot e^k] = P[e_{k+1} \mid h_j \cdot b \cdot c_{k+1}]$

$\therefore \quad P[e^n \mid h_j \cdot b \cdot c^n] = \prod_{k=1}^{n} P[e_k \mid h_j \cdot b \cdot c_k]$

Definition: o_{ku} is a **falsifying outcome** of c_k for h_j with respect to h_i iff

$$P[o_{ku} \mid h_j \cdot b \cdot c_k] = 0 \text{ but } P[o_{ku} \mid h_i \cdot b \cdot c_k] > 0$$

Definition: h_j is **outcome-compatible** with h_i on c_k iff none of the outcomes of c_k are **falsifying** for h_j with respect to h_i

Theorem 1: The Falsification Theorem:
Suppose c^n contains a sub-sequence consisting of m experiments or observations such that for each of them the likelihood of obtaining a *falsifying outcome* is no less than some number $\delta > 0$

i.e., $P[\vee\{o_{ku} : P[o_{ku} \mid h_j \cdot b \cdot c_k] = 0\} \mid h_i \cdot b \cdot c_k] \geq \delta$.

Then,

$$P[\vee\{e^n : P[e^n \mid h_j \cdot b \cdot c^n]/P[e^n \mid h_i \cdot b \cdot c^n] = 0\} \mid h_i \cdot b \cdot c^n]$$

$$\geq 1-(1-\delta)^m.$$

(Notice: if there is a *crucial experiment* in evidence stream c^n, then we may choose $m = 1$ and $\delta = 1$.)

A measure of the Empirical Distinctness of Hypotheses when Outcome-Compatible on the experiment

Definitions: Quality of Information from an outcome

$$\mathrm{QI}[o_{ku} \mid h_i/h_j \mid b \cdot c_k] = \log(P[o_{ku} \mid h_i \cdot b \cdot c_k] \,/\, P[o_{ku} \mid h_j \cdot b \cdot c_k])$$
$$= \log(P[o_{ku} \mid h_i \cdot b \cdot c_k]) - \log(P[o_{ku} \mid h_j \cdot b \cdot c_k])$$

$$\mathrm{QI}[e^n \mid h_i/h_j \mid b \cdot c^n] = \log(P[e^n \mid h_i \cdot b \cdot c^n] \,/\, P[e^n \mid h_j \cdot b \cdot c^n])$$

$$\therefore \quad \mathrm{QI}[e^n \mid h_i/h_j \mid b \cdot c^n] = \sum_{k=1}^{n} \mathrm{QI}[e_k \mid h_i/h_j \mid b \cdot c_k]$$

Definition: Expected Quality of information for an observation or experiment:

for c_k on which h_j is outcome-compatible with h_i,

$EQI[c_k \mid h_i/h_j \mid h_i \cdot b] = \sum_u \; QI[o_{ku} \mid h_i/h_j \mid b \cdot c_k] \cdot P[o_{ku} \mid h_i \cdot b \cdot c_k]$

$EQI[c^n \mid h_i/h_j \mid h_i \cdot b] = \sum_{e^n} QI[e^n \mid h_i/h_j \mid b \cdot c^n] \cdot P[e^n \mid h_i \cdot b \cdot c^n]$

$\therefore \quad EQI[c^n \mid h_j/h_i \mid h_i \cdot b] = \sum_{k=1}^{n} EQI[c_k \mid h_j/h_i \mid h_i \cdot b]$

Definition: $\underline{EQI}[c^n \mid h_i/h_j \mid h_i \cdot b] = EQI[c^n \mid h_i/h_j \mid h_i \cdot b] \div n$

Theorem: Boundedness of EQI

$EQI[c_k \mid h_j/h_i \mid h_i \cdot b] \geq 0$; and

$EQI[c_k \mid h_j/h_i \mid h_i \cdot b] > 0$

if and only if

for at least one of its possible outcomes o_{ku},

$P[o_{ku} \mid h_i \cdot b \cdot c_k] \neq P[o_{ku} \mid h_j \cdot b \cdot c_k]$.

Definition: Variance in the Quality of Information for c_k:

for c_k on which h_j is outcome-compatible with h_i,

$\mathrm{VQI}[c_k \mid h_i/h_j \mid h_i \cdot b] =$

$\sum_u \; (\mathrm{QI}[o_{ku} \mid h_i/h_j \mid b \cdot c_k] - \mathrm{EQI}[c_k \mid h_i/h_j \mid h_i \cdot b])^2 \cdot P[o_{ku} \mid h_i \cdot b \cdot c_k]$

$\mathrm{VQI}[c^n \mid h_i/h_j \mid h_i \cdot b] =$

$\sum_{e^n} (\mathrm{QI}[e^n \mid h_i/h_j \mid b \cdot c^n] - \mathrm{EQI}[c^n \mid h_i/h_j \mid h_i \cdot b])^2 \cdot P[e^n \mid h_i \cdot b \cdot c^n]$

$\therefore \quad \mathrm{VQI}[c^n \mid h_i/h_j \mid h_i \cdot b] = \sum_{k=1}^{n} \mathrm{VQI}[c_k \mid h_i/h_j \mid h_i \cdot b]$

Definition: $\underline{\mathrm{VQI}}[c^n \mid h_i/h_j \mid h_i \cdot b] = \mathrm{VQI}[c^n \mid h_i/h_j \mid h_i \cdot b] \div n$

Theorem 2:
Non-falsifying Likelihood Ratio Convergence Theorem

Choose positive $\varepsilon < 1$, as small as you like, but large enough that (for the number of observations n being contemplated) the value of $\underline{\mathrm{EQI}}[c^n \mid h_i/h_j \mid h_i \cdot b] > -(\log \varepsilon)/n$. Then

$$P[\vee\{e^n : P[e^n \mid h_j \cdot b \cdot c^n]/P[e^n \mid h_i \cdot b \cdot c^n] < \varepsilon\} \mid h_i \cdot b \cdot c^n] \geq$$

$$1 - \frac{1}{n} \cdot \frac{\underline{\mathrm{VQI}}[c^n \mid h_i/h_j \mid h_i \cdot b]}{[\underline{\mathrm{EQI}}[c^n \mid h_i/h_j \mid h_i \cdot b] + (\log \varepsilon)/n]^2}$$

Theorem 2*:
Non-falsifying Likelihood Ratio Convergence Theorem

Suppose there is some fraction γ, $0 < \gamma \leq (1/e)^2$ ($\approx .135$) such that for each possible outcome o_{ku} of each observation condition c_k in c^n, either $P[o_{ku} \mid h_i \cdot b \cdot c_k] = 0$ or $P[o_{ku} \mid h_j \cdot b \cdot c_k] / P[o_{ku} \mid h_i \cdot b \cdot c_k] \geq \gamma$.

Choose positive $\varepsilon < 1$ such that $\underline{EQI}[c^n \mid h_i/h_j \mid h_i \cdot b] > -(\log \varepsilon)/n$. Then

$$P[\vee\{e^n : P[e^n \mid h_j \cdot b \cdot c^n]/P[e^n \mid h_i \cdot b \cdot c^n] < \varepsilon\} \mid h_i \cdot b \cdot c^n] \geq$$

$$1 - \frac{1}{n} \cdot \frac{(\log \gamma)^2}{[\underline{EQI}[c^n \mid h_i/h_j \mid h_i \cdot b] + (\log \varepsilon)/n]^2}$$

Directional Agreement Condition: For each experiment or observation c and each of its possible outcomes o, the *likelihood ratios agree in direction*: i.e.,

$P_\alpha[o \mid h_j \cdot b \cdot c] \,/\, P_\alpha[o \mid h_i \cdot b \cdot c] > 1$ iff

$P_\beta[o \mid h_j \cdot b \cdot c] \,/\, P_\beta[o \mid h_i \cdot b \cdot c] > 1$, and

$P_\alpha[o \mid h_j \cdot b \cdot c] \,/\, P_\alpha[o \mid h_i \cdot b \cdot c] < 1$ iff

$P_\beta[o \mid h_j \cdot b \cdot c] \,/\, P_\beta[o \mid h_i \cdot b \cdot c] < 1$.