The 6 boxes toy model

Alberto Garfagnini

Università di Padova

AA 2022/2023 - Stat Lect. 4

The 6 Boxes Sampling Experiment

The Game

- 6 indistinguishable boxes are prepared with 5 black & white stone
- the composition differs for each box
- boxes are labeled H_j, according to the numbers of white stones in the box, with $j = 0, 1, \dots, 5$

H_0	H_1	H_2	H_3	H_4	H_5
●●●●●	○●●●●	○○●●●	○○○●●	○○○○●	○○○○○

The Rules of the Game

- we choose one box, randomly
- we try to infer the box content (i.e. the box id) by extracting at random one stone from the box
- the extracted stone is reinserted in the box (sampling with replacement)

The 6 Boxes Sampling Experiment

Our Background Information, I

- the following propositions are defined :

  H_j : box j is selected ($j = 0, 1, \ldots, 5$)

  E_w : a white stone is extracted

  E_b : a black stone is extracted

Our Quests

1) what is the probability of selecting one box ?

2) with the extraction of one stone, what is the probability of observing white, $P(E_w|I)$, or black, $P(E_b|I)$ on the next draw ?

3) how does the probability of the next extraction changes after the stone is extracted, and its color known ?

The space Ω of the events

- the following relations apply:

$$\bigcup_{j=0}^{5} H_j = \Omega, \quad \text{and} \quad \bigcup_{k=b}^{w} E_k = \Omega$$

- in general, we are uncertain about all the combinations of E_k and H_j: the 12 constituents, $E_k \cap H_j$ do not share the same probability
- as an example:

$$P(E_w \cdot H_0 \big| I) = 0, \quad P(E_w \cdot H_5 \big| I) = 1$$

- E_k and H_j form a complete class of hypotheses, each event can be written as a logical sum of the constituents:

$$E_k = \bigcup_j \left(E_k \bigcap H_j \right), \quad \text{and} \quad H_j = \bigcup_k \left(E_k \bigcap H_j \right)$$

- since the events $E_k \cap H_j$ are mutually exclusive, by construction, we have:

$$P(E_k) = \sum_j P(E_k \cdot H_j \big| I) = \sum_j P(E_k \big| H_j I)\, P(H_j \big| I)$$

- and

$$P(H_j) = \sum_k P(H_j \cdot E_k \big| I) = \sum_k P(H_j \big| E_k I)\, P(E_k \big| I)$$

The Process of Knowledge

- E_k is an observable effect: we can experience it with our senses
- H_j is a physical hypothesis: it is not directly observable

  Another rule of the game: we are not allowed to look inside the box !

  ➜ H_j are the possible causes of the effect
- Inference : guessing the causes from the effects

Our experiment consists in

1. extracting stones, randomly and with replacement, from an unknown box
2. evaluating the probability that the box is one of the six boxes

➜ aim of each measurement: update our beliefs about each cause, given all available information

and our calculations

- after the first extraction, $E^{(1)}$, we will compute:

$$P\left(H_j \,\middle|\, E^{(1)} I\right)$$

- and, after the second extraction $E^{(2)}$:

$$P\left(H_j \,\middle|\, E^{(1)} E^{(2)} I\right)$$

- and so forth
- what can be easily calculated is the probability of observing the different effects, giving each cause, $P(E_k \,|\, H_j I)$:

$$P\left(E_w \,\middle|\, H_j I\right) = \frac{j}{5}, \quad \text{and} \quad P\left(E_b \middle| H_j I\right) = 1 - P\left(E_w \,\middle|\, H_j I\right) = \frac{5-j}{5}$$

and our calculations ...

- the product rule

$$P\left(E_k H_j \middle| I\right) = P\left(E_k \middle| H_j I\right) P\left(H_j \middle| I\right)$$
$$= P\left(H_j \middle| E_k I\right) P\left(E_k \middle| I\right)$$

- can be rewritten as

$$\frac{P\left(E_k \middle| H_j I\right)}{P\left(E_k \middle| I\right)} = \frac{P\left(H_j \middle| E_k I\right)}{P\left(H_j \middle| I\right)}$$

- we know $P(E_k \middle| H_j I)$ and $P(E_k|I)$ can be evaluated as:

$$P\left(E_k \middle| I\right) = \sum_j P\left(E_k \middle| H_j I\right) P\left(H_j \middle| I\right) = \frac{0+1+2+3+4+5}{5} \cdot \frac{1}{6} = \frac{1}{2}$$

- as we would expect

and our calculations ... ...

- we can rewrite the product rule as

$$\frac{P\left(H_j \middle| E_k I\right)}{P\left(H_j \middle| I\right)} = \frac{P\left(E_k \middle| H_j I\right)}{P\left(E_k \middle| I\right)} = 2 \cdot P\left(E_k \middle| H_j I\right)$$

- in case of a white stone, $P(E_w \middle| I) = 1$,

$$\frac{P\left(H_j \middle| E_w I\right)}{P\left(H_j \middle| I\right)} = 2 \cdot \frac{j}{5}$$

- while, for a black stone, $P(E_b \middle| I) = 1$,

$$\frac{P\left(H_j \middle| E_b I\right)}{P\left(H_j \middle| I\right)} = 2 \cdot \frac{5-j}{5}$$

and our calculations ... ...

- putting all the ingredients together, we get Bayes' theorem

$$P\left(H_j \mid E_k I \right) = \frac{P\left(E_k \mid H_j I \right) P\left(H_j \mid I \right)}{\sum_j P\left(E_k \mid H_j I \right) P\left(H_j | I \right)}$$

- the denominator is just a normalization factor, and we can simply write:

$$P\left(H_j \middle| E_k I \right) \propto P\left(E_k \middle| H_j I \right) P\left(H_j \middle| I \right)$$

- or, in clear text

$$\text{Posterior} \propto \text{Likelihood} \times \text{Prior}$$

- Bayes' theorem is simply a compact representation of what has been done in the previous steps.
- it is a formal tool for updating beliefs using logic instead of only intuition

Running the experiment

- we randomly select a box, and start to sample stones from the box
- after each extraction, we update the probabilities of each hypothesis, using Bayes' theorem:

$$P\left(H_j \middle| I_n \right) = \frac{P\left(E^{(n)} \middle| H_j\, I_{n-1}\right) P\left(H_j \middle| I_{n-1}\right)}{\sum_l P\left(E^{(n)} \middle| H_l\, I_{n-1}\right) P\left(H_l \middle| I_{n-1}\right)}$$

- where $E^{(n)}$ refers to the n-th extraction,
- $P(E^{(n)} | H_j)$ have been computed before:

$$P(E_w^{(n)} | H_j) = \frac{j}{5}, \quad P(E_b^{(n)} | H_j) = \frac{5-j}{5}$$

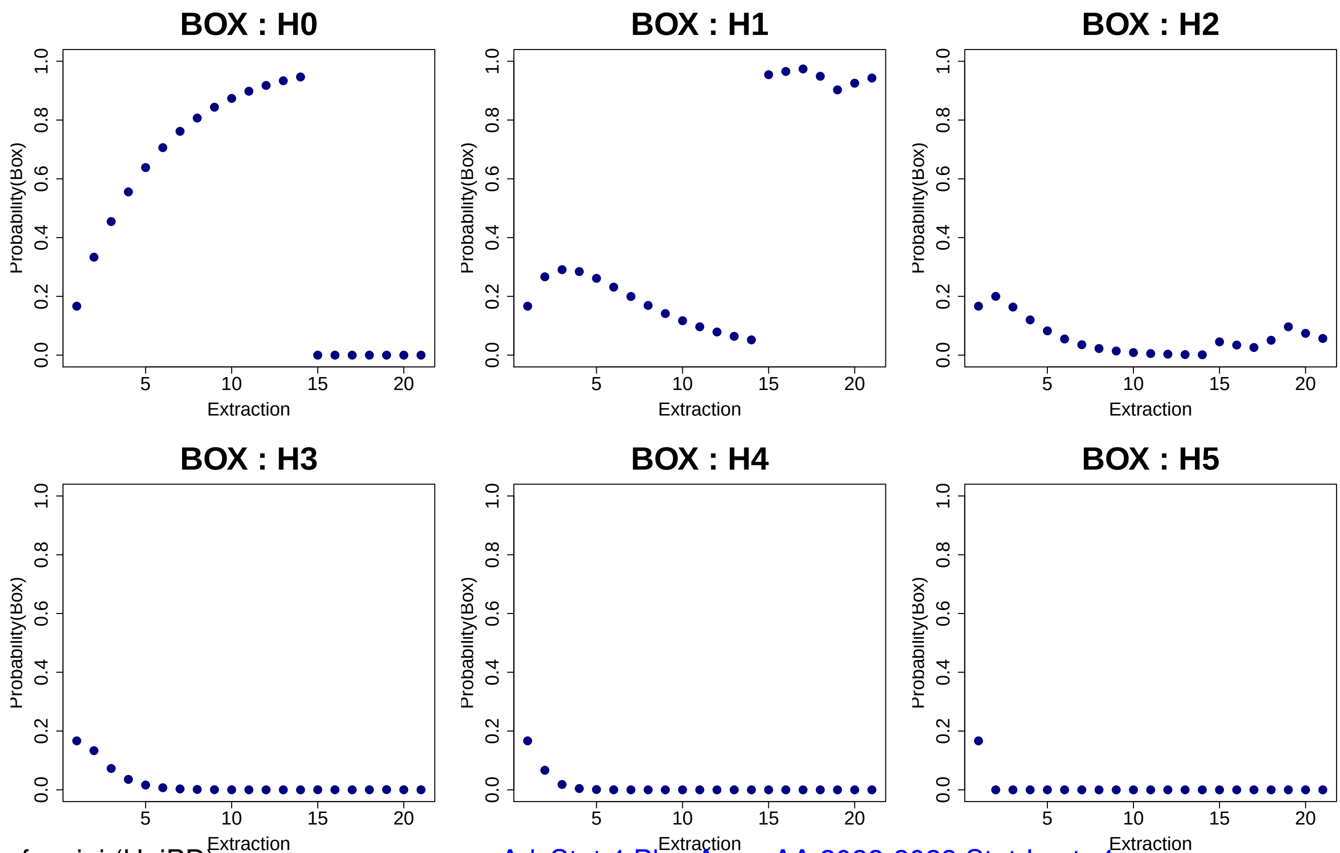
- and $P(H_j | I_{n-1})$ have been given by the calculations at extraction $(n-1)$-th

Running the experiment

Trial	E	H_0	H_1	H_2	H_3	H_4	H_5	$P(E_w \| I_n)$
0	-	0.167	0.167	0.167	0.167	0.167	0.167	0.5
1	B	0.33	0.27	0.2	0.13	0.06	0	0.27
2	B	0.45	0.29	0.163	0.073	0.0182	0	0.18
3	B	0.55	0.28	0.12	0.036	0.004	0	0.13
4	B	0.64	0.26	0.08	0.016	0.001	0	0.096
5	B	0.71	0.23	0.05	0.007	2.2E-4	0	0.072
6	B	0.76	0.20	0.04	0.003	4.9e-5	0	0.056
7	B	0.81	0.17	0.02	0.001	1.0e-5	0	0.044
8	B	0.84	0.14	0.01	5.5e-4	2.2e-6	0	0.034
9	B	0.87	0.12	0.009	2.3e-4	4.5e-7	0	0.027
10	B	0.90	0.10	0.005	9.4e-5	9.2e-8	0	0.022
11	B	0.92	0.08	0.003	3.8e-5	1.9e-8	0	0.017
12	B	0.93	0.06	0.002	1.6e-5	3.8e-9	0	0.014
13	B	0.95	0.05	0.001	6.3e-6	7.8e-10	0	0.011
14	W	0	0.95	0.045	3.5e-4	5.7e-8	0	0.21
20	B	0	0.93	7.4e-2	3.8e-4	1.4e-8	0	0.21
40	W	0	0.998	1.4e-3	7.1e-9	8.7e-19	0	0.20

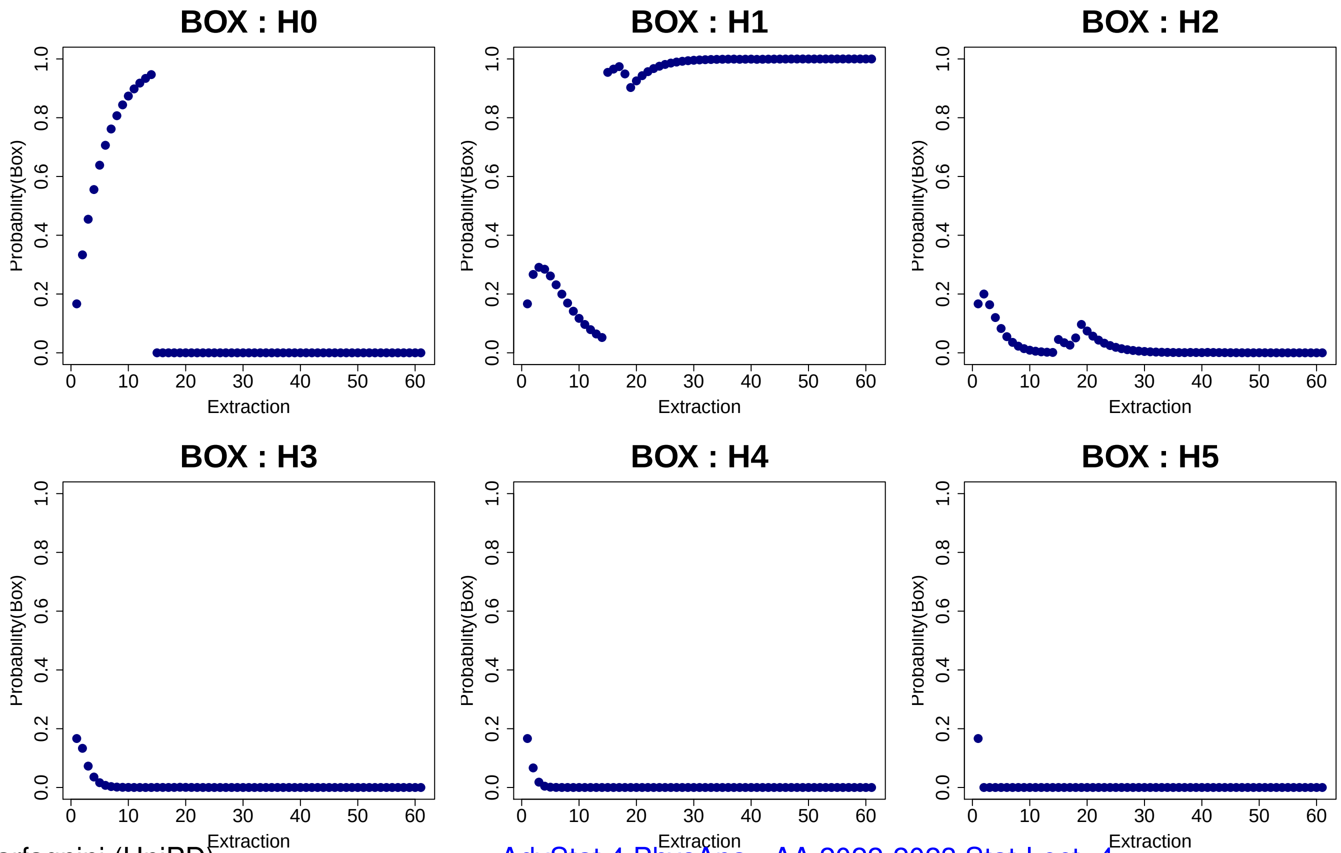
Run results : 20 samplings

- Run performed with `set.seed(89540)`
- important extraction at round 14



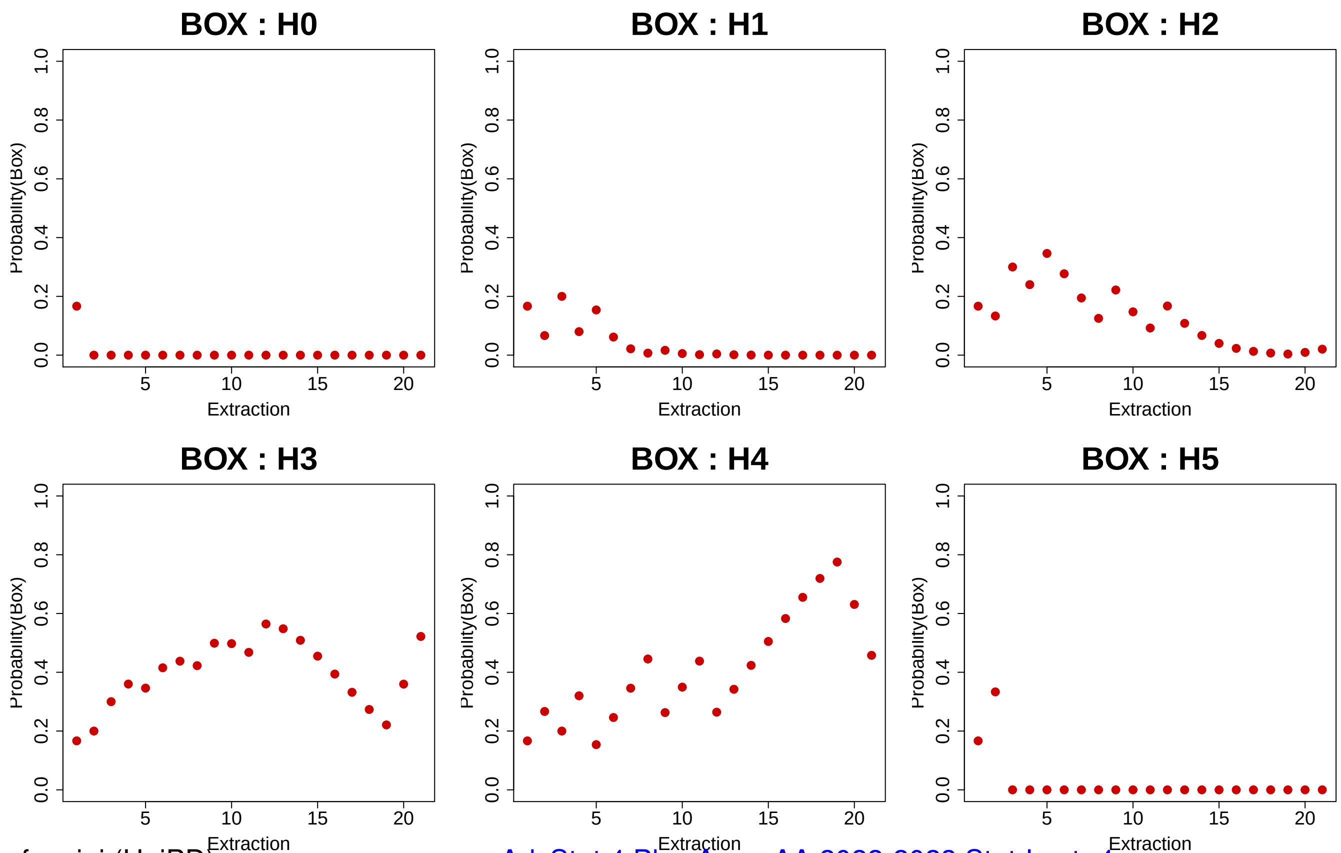
Run results : 60 samplings

- Box H_1 is the most probable : $\boxed{\circ\bullet\bullet\bullet\bullet}$ $P(E_w|I_n) = 0.2$, as expected



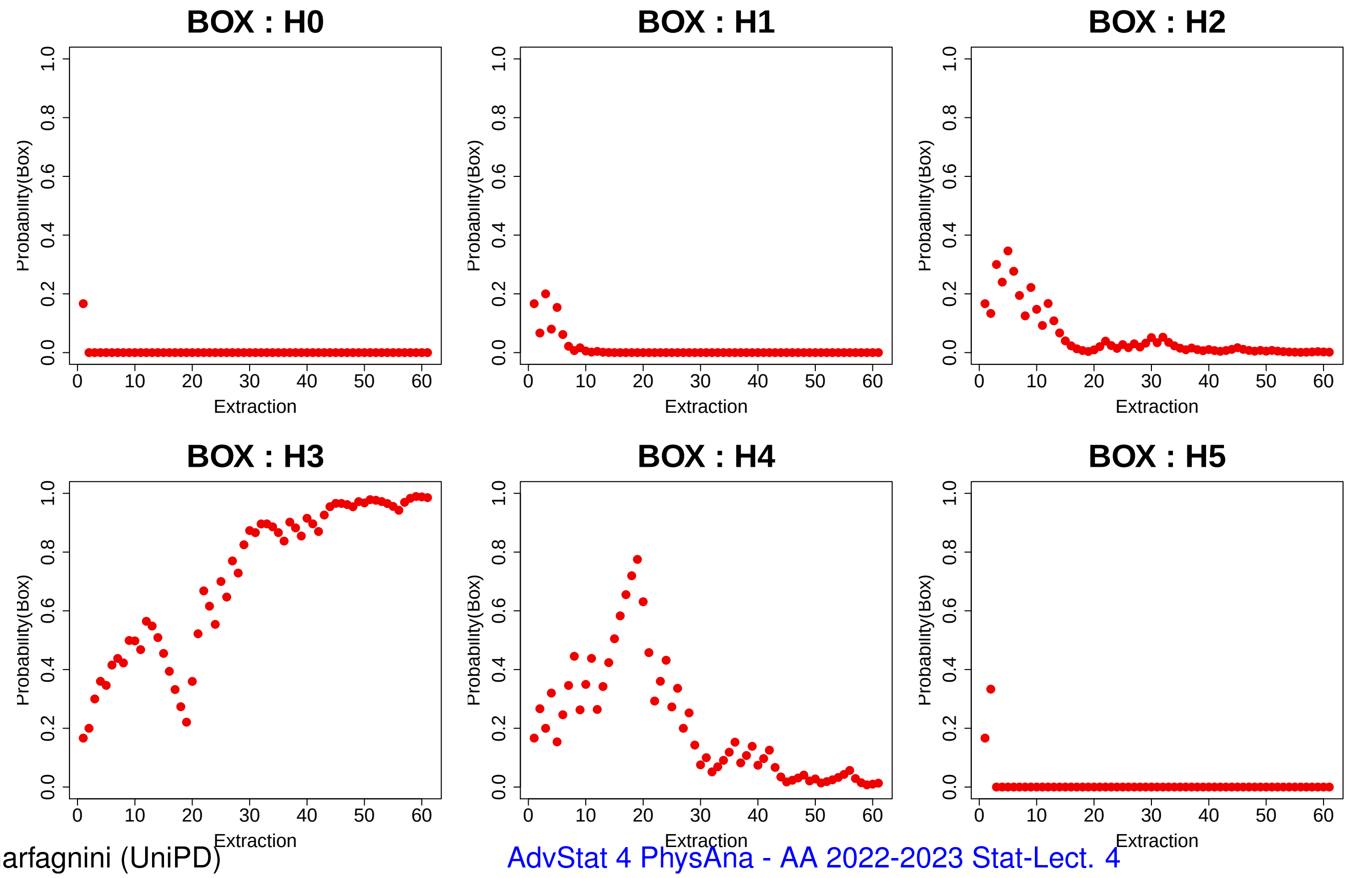
New run results: 20 samplings

- Run performed with `set.seed(89540)`
- most flavored oscillates between H_3 and H_4



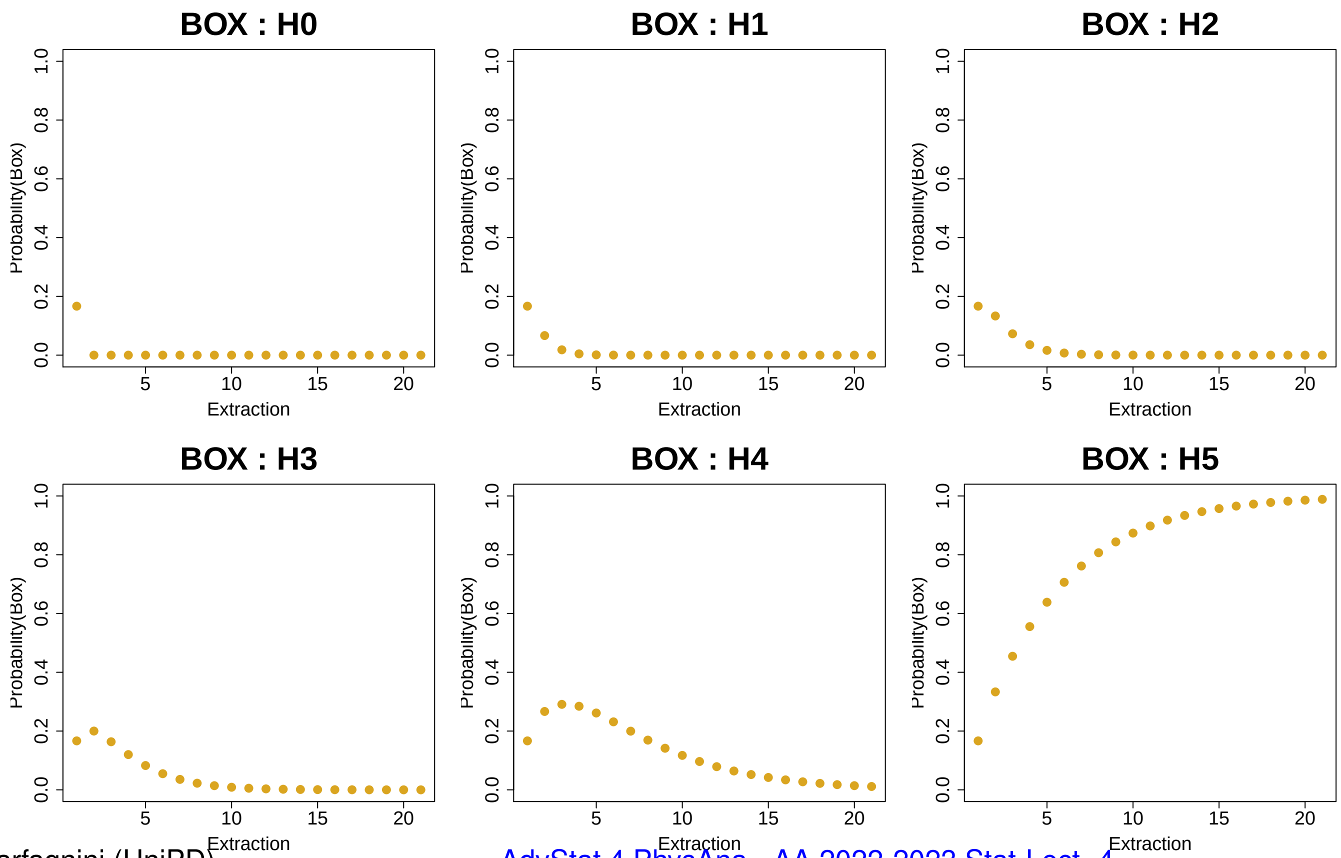
New run results : 60 samplings

- Box H_3 is the most probable : $\boxed{\circ\circ\circ\bullet\bullet}$ $P(E_w|I_n) = 0.6$, as expected



Run with an extreme box

- Run performed with `set.seed(89540)` and box $\boxed{\circ\circ\circ\circ\circ}$



References for the 6 Boxes Toy Model

Articles

- G. D'Agostini,*Teaching statistics in the physics curriculum: Unifying and clarifying role of subjective probability*, Am. Jour. Phys. 67, 1260 (1999), arXiv:physics/9908014

- G. D'Agostini,*More lessons from the six box toy experiment*, arXiv:1701.01143

- G. D'Agostini, *Probability, propensity and probabilities of propensities (and of probabilities)*, arXiv:1612.05292

Additional Material

- G. D'Agostini Web Page at University of Rome, La Sapienza, `http://www.roma1.infn.it/~dagos/teaching.html`