Comments on Jan-Willem Romeijn's *Abducted by Bayesians?*

Franz Huber, California Institute of Technology

In his paper "Abducted by Bayesians?" Jan-Willem Romeijn assumes that any difference between hypotheses with equal likelihoods is theoretical. Romeijn then argues that "within Bayesian statistical inference, such notions can play an active role" – i.e. two empirically equivalent hypotheses can receive different empirical support. This leads Romeijn to the two further claims that (a) underdetermination should be seen as a methodological tool, and that (b) the use of theoretical notions in Bayesian statistical inference may be viewed as an explication of abduction.

I buy Romeijn's assumption, but not his conclusion. Hence, I claim that his argument is not valid. More specifically, I will argue that Romeijn does not succeed in showing that two empirically equivalent hypotheses can receive different empirical support; and even if he did, his explication of abduction is heading in the wrong direction.

1 Empirical Equivalence?

I will restrict the discussion to a discrete version of Romeijn's example. We are considering a series of coin tosses. Q_i^0 and Q_i^1, respectively, are the propositions that the ith toss yields heads and tails, respectively. The background information includes the assumption that the coin tosses Q_i are independent and identically distributed. We also have the following partition of hypotheses: $H_0, H_{1/2}, H_1$, where H_θ says that the chance that the coin lands tails (on any toss) equals θ. E_t is a sequence of t observations that specifies for each toss $i \in \{1, \ldots, t\}$ whether the coin lands heads, Q_i^0, or tails, Q_i^1. Using Lewis' Principal Principle, including the admissibility of historical information, Romeijn infers that

$$\Pr\left(Q_{t+1}^1 \mid H_\theta \cap E_t\right) = \theta, \quad \theta = 0, 1/2, 1$$

Together with the prior probabilities of the H_θ, $\Pr(H_\theta)$, Bayes' Theorem gives us their posterior probabilities, e.g.

$$\Pr\left(H_{1/2} \mid E_t\right) = \frac{\Pr\left(E_t \mid H_{1/2}\right) \cdot \Pr\left(H_{1/2}\right)}{\sum_\theta \Pr\left(E_t \mid H_\theta\right) \cdot \Pr\left(H_\theta\right)}$$

Romeijn then focuses on "inductive inferences using two duplicate subpartitions, which differ only in the entirely theoretical property that they posit different mechanisms underlying the observations." The difference in the underlying mechanism is expressed by two different assignments of prior probabilities to the H_θ. Romeijn then claims that "because these priors react to the updating operations differently, the partitions can be distinguished by the observations after all, even while they consist of statistically identical hypotheses." I claim that the two partitions are not empirically equivalent. No wonder they receive different empirical support.

My discrete version of Romeijn's example gives us

$$\Pr{}_0(H_0) = \Pr{}_0(H_1) = \varepsilon, \quad \Pr{}_0\left(H_{1/2}\right) = 1 - 2\varepsilon$$

for the coin form the ordinary wallet, and

$$\Pr{}_1(H_0) = \Pr{}_1(H_1) = 1/2 - \varepsilon, \quad \Pr{}_1\left(H_{1/2}\right) = 2\varepsilon$$

for the coin from the conjurer's box. Then we weigh these two probability measures as follows: $\Pr(\Pr_0) = \Pr(\Pr_1) = 1/2$. This gives us

$$\Pr(H_\theta) = \Pr{}_0(H_\theta) \cdot \Pr(\Pr{}_0) + \Pr{}_1(H_\theta) \cdot \Pr(\Pr{}_1)$$

(As an aside, it seems more appropriate to model Romeijn's scenario by a single Pr with $\Pr(H_0) = \Pr(H_1) = 1/4$ and $\Pr\left(H_{1/2}\right) = 1/2$. For what is the interpretation of $\Pr(\Pr_1(H_\theta))$? Is it the agent's actual credence that an ideal agent would assign $\Pr_1(H_\theta)$ to the proposition that H_θ is true?) If we observe a sequence of six coin tosses all landing tails, $E_6 = Q_1^1 \cap \ldots \cap Q_6^1$, we find that

$$\begin{aligned}
\Pr(\Pr{}_0 \mid E_6) &= \frac{\Pr(E_6 \mid \Pr{}_0) \cdot \Pr(\Pr{}_0)}{\Pr(E_6 \mid \Pr{}_0) \cdot \Pr(\Pr{}_0) + \Pr(E_6 \mid \Pr{}_1) \cdot \Pr(\Pr{}_1)} \\
&= \frac{\Pr(E_6 \mid \Pr{}_0)}{\Pr(E_6 \mid \Pr{}_0) + \Pr(E_6 \mid \Pr{}_1)}
\end{aligned}$$

$$= \frac{\Pr(E_6 \mid H_0) \cdot \Pr{}_0(H_0) + \Pr\left(E_6 \mid H_{1/2}\right) \cdot \Pr{}_0\left(H_{1/2}\right) + \Pr(E_6 \mid H_1) \cdot \Pr{}_0(H_1)}{\Pr(E_6 \mid H_0) \cdot (\Pr{}_0(H_0) + \Pr{}_1(H_0)) + \Pr\left(E_6 \mid H_{1/2}\right) \cdot \left(\Pr{}_0\left(H_{1/2}\right) + \Pr{}_1\left(H_{1/2}\right)\right) + \Pr(E_6 \mid H_1) \cdot (\Pr{}_0(H_1) + \Pr{}_1(H_1))}$$

$$
\begin{aligned}
&= \frac{0^6 \cdot \varepsilon + 1/2^6 \cdot (1-2\varepsilon) + 1^6 \cdot \varepsilon}{0^6 \cdot (\varepsilon + 1/2 - \varepsilon) + 1/2^6 \cdot (1 - 2\varepsilon + 2\varepsilon) + 1^6 \cdot (\varepsilon + 1/2 - \varepsilon)} \\
&= \frac{1/2^6 \cdot (1-2\varepsilon) + \varepsilon}{1/2^6 + 1/2} \\
&= \frac{1/64 + 62\varepsilon/64}{33/64} \\
&= \frac{1+62\varepsilon}{33} \\
&< 1/2 \\
&= \Pr(\Pr_0)
\end{aligned}
$$

$$
\begin{aligned}
\Pr(\Pr_1 \mid E_6) &= \frac{\Pr(E_6 \mid \Pr_1) \cdot \Pr(\Pr_1)}{\Pr(E_6 \mid \Pr_1) \cdot \Pr(\Pr_1) + \Pr(E_6 \mid \Pr_0) \cdot \Pr(\Pr_0)} \\
&= \frac{\Pr(E_6 \mid \Pr_1)}{\Pr(E_6 \mid \Pr_1) + \Pr(E_6 \mid \Pr_0)}
\end{aligned}
$$

$$
= \frac{\Pr(E_6 \mid H_0) \cdot \Pr_1(H_0) + \Pr\left(E_6 \mid H_{1/2}\right) \cdot \Pr_1\left(H_{1/2}\right) + \Pr(E_6 \mid H_1) \cdot \Pr_1(H_1)}{\Pr(E_6 \mid H_0) \cdot (\Pr_1(H_0) + \Pr_0(H_0)) + \Pr\left(E_6 \mid H_{1/2}\right) \cdot \left(\Pr_1\left(H_{1/2}\right) + \Pr_0\left(H_{1/2}\right)\right) + \Pr(E_6 \mid H_1) \cdot (\Pr_1(H_1) + \Pr_0(H_1))}
$$

$$
\begin{aligned}
&= \frac{0^6 \cdot (1/2 - \varepsilon) + 1/2^6 \cdot 2\varepsilon + 1^6 \cdot (1/2 - \varepsilon)}{0^6 \cdot (1/2 - \varepsilon + \varepsilon) + 1/2^6 \cdot (2\varepsilon + 1 - 2\varepsilon) + 1^6 \cdot (1/2 - \varepsilon + \varepsilon)} \\
&= \frac{1/2^6 \cdot 2\varepsilon + 1/2 - \varepsilon}{1/2^6 + 1/2} \\
&= \frac{32/64 - 62\varepsilon/64}{33/64} \\
&= \frac{32 - 62\varepsilon}{33} \\
&> 1/2 \\
&= \Pr(\Pr_1)
\end{aligned}
$$

As Romeijn puts it, "the two subpartitions [...] are observationally indistinguishable, but the different expectations over these partitions [...] make the partitions observationally distinct after all. In other words, as a side effect of updating over the separate subpartitions, the observations become relevant to the theoretical distinction between hypotheses [$\Pr_0$] and [$\Pr_1$]."

However, as is clear from the above calculation, the two "hypotheses" Pr_0 and Pr_1 are not observationally indistinguishable by Romeijn's own lights. Reason being, they have different likelihoods. More precisely, although both "hypotheses" say that the coin tosses are independent and identically distributed and such that

$$\Pr\left(Q_i^1 \mid \Pr{}_0\right) = \Pr\left(Q_i^1 \mid \Pr{}_1\right) = 1/2,$$

not all likelihoods are the same. For instance, for $E_6 = Q_1^1 \cap \ldots \cap Q_6^1$ we get

$$\begin{aligned}
\Pr\left(E_6 \mid \Pr{}_0\right) &= \Pr\left(E_6 \mid H_0\right)\Pr{}_0\left(H_0\right) + \Pr\left(E_6 \mid H_{1/2}\right)\Pr{}_0\left(H_{1/2}\right) + \Pr\left(E_6 \mid H_1\right)\Pr{}_0\left(H_1\right) \\
&= 0^6 \cdot \varepsilon + 1/2^6 \cdot (1-2\varepsilon) + 1^6 \cdot \varepsilon \\
&= 1/64 + 62\varepsilon/64 \\
\Pr\left(E_6 \mid \Pr{}_1\right) &= \Pr\left(E_6 \mid H_0\right)\Pr{}_1\left(H_0\right) + \Pr\left(E_6 \mid H_{1/2}\right)\Pr{}_1\left(H_{1/2}\right) + \Pr\left(E_6 \mid H_1\right)\Pr{}_1\left(H_1\right) \\
&= 0^6 \cdot (1/2-\varepsilon) + 1/2^6 \cdot 2\varepsilon + 1^6 \cdot (1/2-\varepsilon) \\
&= 32/64 - 62\varepsilon/64
\end{aligned}$$

Hence the two "hypotheses" Pr_0 and Pr_1 are not empirically equivalent to begin with, and it is thus not surprising that they receive different empirical support.

Furthermore, what Romeijn calls a "hypothesis" is a prior probability over a partition. The reason he gives for calling two priors over the same partition empirically equivalent is that the two (sub)partitions consist of the same hypotheses. But if that were sufficient for empirical equivalence, *any* two hypotheses – both priors over partitions as well as ordinary propositions – are or can be made empirically equivalent.

Any two priors over the same partition are empirically equivalent. If we have two priors Pr_0 and Pr_1, respectively, over two distinct partitions $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively, then one can reformulate these as priors over the partition $\mathcal{H} = \{H \cap H' : H \in \mathcal{H}_0, H' \in \mathcal{H}_1\}$. Then we could define $\mathrm{Pr}_0\left(H \cap H'\right) = \mathrm{Pr}_0\left(H\right)$, where H' is the first element (in some enumeration of the elements) of $\mathcal{H}_1$ such that $H \cap H'$ is not empty, and $\mathrm{Pr}_0\left(H \cap H''\right) = 0$ for the remaining elements H'' of $\mathcal{H}_1$; similarly for Pr_1. Ordinary propositions H can finally be made empirically equivalent by identifying them with the prior over the partition $\left\{H, \overline{H}\right\}$ such that $\Pr\left(H\right) = 1$. To be sure, I do, of course, not want to argue for this identification. But it is an all too natural move if priors over partitions were indeed hypotheses.

2 Abduction?

In the second part of his paper Romeijn uses his example to draw conclusions for the methodology of science in general. “The second claim follows from the observation that in cases like the above one, Bayesian statistics provides the reasons for choosing between hypotheses that are only theoretically distinct. Such choices are usually considered to involve abductive inference. We may therefore argue that the above cases provide a first sketch of a Bayesian model of abductive inference”.

As I argued in the previous section, Romeijn still has to provide an example of two empirically equivalent hypotheses that are treated differently in an interesting way. But then, he could make his points concerning abduction independently of his claims concerning the use of purely theoretical distinctions. For contrary to what Romeijn suggests, the purpose of abduction is not to “enable us to choose between theoretical superstructures on the basis of explanatory considerations or other theoretical virtues.” An abductive inference, even though referring to explanatory considerations, does not necessarily refer to theoretical virtues.

In either case, what Romeijn seems to have in mind for his explication of abduction are cases where one hypothesis has a higher posterior than another, even though the two hypotheses are empirically equivalent, and hence have the same likelihoods. The underlying abductive inference rule thus seems to be that one should infer those hypotheses that maximize posterior probability. Now it might make sense for a Bayesian to say that abduction consists in maximizing likelihoods (the idea being that likelihoods measure explanatory power), or that abduction consists in maximizing incremental confirmation alias expected informativeness. But if anything is clear about abduction, it is that it does not aim at the most probable hypothesis, and thus does not consist in maximizing posterior probability. For the observational content of any theoretical hypothesis, if finitely axiomatizable, always receives a posterior that is at least as high, of not higher, than that of the hypothesis in question. Romeijn’s sketch of an explication is thus heading in the wrong direction.