Random Emeralds*

Sven Neth

Abstract

Suppose we observe many emeralds which are all green. This observation usually provides good evidence that all emeralds are green. However, the emeralds we have observed are also all *grue*, which means that they are either green and already observed or blue and not yet observed. We usually do not think that our observation provides good evidence that all emeralds are grue. Why? I argue that if we are in the best case for inductive reasoning, we have reason to assign low probability to the hypothesis that all emeralds are grue before seeing any evidence. My argument appeals to random sampling and the observation-independence of green, understood as probabilistic independence of whether emeralds are green and when they are observed.

1 Introduction

Suppose we observe many emeralds which are all green. Our observation provides good evidence that all emeralds are green. However, the emeralds we have observed are also all *grue*: either green and already observed or blue and not

*Forthcoming in *Philosophical Quarterly*. Thanks to Shamik Dasgupta, Branden Fitelson, Aglaia von Götz, Wesley Holliday, Simon Huttegger, Ethan Jerzak, Hanti Lin, Taylor Madigan, Alex Meehan, Krzysztof Mierzewski, Milan Mossé, Lars Neth, Adrian Ommundsen, Jan-Willem Romeijn, Cat Saint-Croix, Pia Schneider, Brian Skyrms, Jan Sprenger, Jeremy Strasser, Michael Strevens, Mack Sullivan, Michael Titelbaum, Jared Warren and two anonymous referees, to audiences in Chicago, Berkeley and Pittsburgh and to my graduate seminar on grue in Spring 2025. Special thanks to Morgan Connolly, Mathias Böhm, Lara Buchak, Mikayla Kelley and John MacFarlane for their unbounded patience in discussing these ideas. This paper is dedicated to the late Barry Stroud, who first taught me about grue in Fall 2014 at Berkeley. I've been puzzled ever since.

yet observed (Goodman 1955). We usually do not think that our observation provides good evidence that all emeralds are grue. Why?

In a Bayesian framework, how much our evidence supports a hypothesis depends on two factors: The conditional probability of our evidence given the hypothesis, which is called the *likelihood*, and the probability of our hypothesis before observing our evidence, which is called the *prior probability.* Both the hypothesis that all emeralds are green (all-green) and the hypothesis that all emeralds are grue (all-grue) assign the same likelihood to our evidence. Both hypotheses predict that we would observe exactly what we do in fact observe. So if there is any difference between all-green and all-grue, it must concern their prior probability. In their classic textbook on Bayesian reasoning, Urbach and Howson write:

> So theories such as [...] Goodman's grue-variants must, for some reason, have lower prior probabilities. (Urbach and Howson 1993, p. 177)[1]

However, assigning a low prior probability to all-grue might seem like "simply some sort of 'a priori' prejudice against 'grue' hypotheses enshrined in a Bayesian formalism" (Fitelson 2008, p. 631). It would be great to have a reason why we should assign a low prior probability to all-grue. My goal is to provide such a reason.

I begin with some brief reflections on why the New Riddle of Induction strikes us as paradoxical. The answer, I suggest, is because it seems to arise even in the best case for inductive reasoning. Then I explain how we can defuse the paradox. If we are in the best case for inductive reasoning—emeralds are randomly sampled and whether emeralds are green is independent of when they are observed—there is an upper bound on the prior probability of all-grue but

[1] The idea that all-grue does or should have a low prior probability is expressed by Chihara (1981), Horwich (1982), Rosenkrantz (1982), Sober (1994), Norton (2006), Sider (2011), Sprenger and Hartmann (2019) and many others. Lewis (1983, p. 375) writes that "[t]he principles of charity will impute a bias toward believing that things are green rather than grue". Moss (2018, p. 101) suggests that "you may be *a priori* justified in believing that the hypothesis that all emeralds are green is more likely than the hypothesis that all emeralds are grue".

no upper bound on the prior probability of all-green.

I consider the objection that there is a symmetrical 'gruesome' argument for an upper bound on the prior probability of all-green and show that this argument fails. The difference between green and grue is that when randomly sampling, green is *observation-independent* while grue is not, understood as probabilistic independence of whether emeralds are green and when they are observed.

2 Paradox Regained

Here is a hypothesis. The New Riddle seems to arise even in the best case for inductive reasoning. Even if we set aside worries about biased sampling, about our observations affecting the system we are observing and so on, it seems like we need a further reason for why we should prefer the all-green hypothesis over the all-grue hypothesis. And it can be difficult to see what this reason could be.

This hypothesis explains the feeling of paradox. In many cases, for example when running clinical trials, everyone would likely agree that inductive reasoning is hard, that we should not be overconfident in our predictions and so on. But since the New Riddle arises even in the best case for inductive reasoning, it threatens the very possibility of inductive reasoning. While Goodman originally formulated the New Riddle to argue against a specific non-probabilistic theory of confirmation due to Hempel (1943), the problem is much more general and arises on a Bayesian theory of confirmation as well.[2]

When responding to the New Riddle, the challenge is to find an asymmetry between green and grue in the best case for inductive reasoning. The asymmetry should be *epistemic* in the sense that it constrains our inductive inferences.[3]

[2] Hooker (1968) and Fitelson (2008) discuss the New Riddle from the perspective of Hempel's confirmation theory and Titelbaum (2022, pp. 215–221) discusses it from the perspective of Carnap's confirmation theory. As Earman (1992, p. 104) observed thirty years ago, "[e]nough ink has been spilled over Goodman's 'new problem of induction' to drown an elephant". The spillage hasn't stopped since. Stalker (1994) and Elgin (1997) provide partial overviews. Neth (ms) provides a very short introduction.

[3] The framing of the New Riddle in terms of an epistemic asymmetry is suggested by Sober (1994). Thomson (1966, p. 300) makes a similar point.

In a Bayesian framework, this means that the asymmetry should constrain our assignment of prior probabilities. Furthermore, the asymmetry should be *language independent* in the sense that it does not depend on our choice of linguistic primitives.

Our hypothesis explains why Goodman and others talk about emeralds. It is natural to assume that the colors of emeralds are independent of our observation. There are different ways to make this intuitive idea precise. Below, I propose to articulate it in terms of probabilistic independence.[4] The example also invites us to set aside worries about biased sampling. To be sure, Goodman never explicitly says that emeralds are randomly sampled and that colors are independent of observation.[5] Rather, these are plausible background assumptions which explain why we feel the force of the paradox.

To bring this out, suppose that emeralds are not randomly sampled but presented to you by a shady broker who also happens to sell bets on the color of the next emerald. In this context, it is not very surprising if we have no reason to expect the next emerald to be green rather than grue. Or suppose we are digging for emeralds and know that green emeralds are lighter than blue emeralds and so tend to be closer to the surface. Again, while it might be difficult to figure out what we should predict about the color of the next emerald, this does not seem paradoxical.

In a nutshell, my response to the New Riddle runs as follows. If we carefully reflect on the background assumptions which characterize the best case for inductive reasoning—random sampling and observation-independent properties—we can see that these assumptions imply an asymmetry between all-green and

[4] Observation-independence can also be understood in counterfactual terms (Jackson 1975; Godfrey-Smith 2002; Warren 2023). While a detailed comparison with this counterfactual response would need more space, the key advantage of using my probabilistic notion is that it connects straightforwardly to the prior. Furthermore, the probabilistic notion of observation-independence is much more closely connected to constraints on priors than the idea that green is 'natural' or 'purely qualitative' (Carnap 1947; Quine 1970; Lewis 1983; Sider 2011).

[5] Although the very first example by Goodman (1946, p. 383) is about drawing marbles from an urn and observing that they are all 'S', which means 'drawn by VE day and red or drawn later and non-red'—a setting which makes it natural to assume random sampling. Hacking (1965, pp. 41–2) also presents (and gets very puzzled by) the New Riddle in the context of random sampling.

all-grue. So if we really are in the best case for inductive reasoning, the apparent symmetry between all-green and all-grue is already broken.

3 The Random Principle

Imagine you are about to observe two emeralds by random sampling. Suddenly, an angel appears before you and tells you that you will observe one green and one blue emerald. You are certain that the angel has spoken truly. Do you have any reason to think you will first observe the green emerald and then the blue emerald rather than the other way around? It seems not, by our assumption of random sampling. So you should assign the same probability to both possible orderings GB and BG. This remains true if you have some other background information consistent with random sampling. For example, you might believe that the emeralds are likely green or likely blue. This background information does not allow you to predict that one of the orderings GB and BG is more likely. More generally, if you are randomly sampling emeralds, you should assign the same probability to any particular ordering of green and blue emeralds given that you will observe a fixed number of green and blue emeralds. This means that your prior should be *exchangeable* with respect to green.

Let me introduce some notation. Suppose we are about to observe some finite number n of emeralds by random sampling. The emeralds are known to be either green or blue. We can model our observations as a sequence of random variables $X_1, X_2, ..., X_n$, where $X_i = 1$ if the i-th emerald is green and $X_i = 0$ otherwise. (You might complain that I'm stacking the deck in favor of all-green by describing everything in green rather than grue language. I will address this worry below.) Intuitively, to say that the sequence $X_1, X_2, ..., X_n$ is exchangeable means that order does not matter.

To make the notion of exchangeability precise, define the *joint distribution* of the sequence of random variables $X_1, ..., X_n$ as $p_{X_1,....,X_n}(x_1, ..., x_n) = \mathbb{P}(X_1 =$

$x_1 \cap ... \cap X_n = x_n$).[6] Now we can define exchangeability as follows:

Definition 1. *The sequence of random variables $X_1, ..., X_n$ is exchangeable iff the joint distribution $p_{X_1,....,X_n}(x_1, ..., x_n)$ is a symmetric function of its arguments (Pitman 1993, p. 238). This means that if two sequences of values are permutations of each other, they are equally likely.*

Consider the sequence of random variables X_1, X_2. If this sequence is exchangeable, we must have $\mathbb{P}(X_1 = 1, X_2 = 0) = \mathbb{P}(X_1 = 0, X_2 = 1)$. Written more succinctly: $\mathbb{P}(10) = \mathbb{P}(01)$. Exchangeability does not constrain what these probabilities are, only that they must be equal. An example of an exchangeable joint distribution is $\mathbb{P}(10) = \mathbb{P}(01) = .5$. Another example of an exchangeable joint distribution is $\mathbb{P}(11) = 1$, since $\mathbb{P}(10) = \mathbb{P}(01) = 0$. An example of a joint distribution which is not exchangeable is $\mathbb{P}(10) = 1$.

Why accept exchangeability? Because it follows from the assumption that we are *randomly sampling* the emeralds and that green is *observation-independent*, which means that when emeralds are observed and whether they are green are probabilistically independent. So here is my argument:

1. We are randomly sampling.

2. Green is observation-independent.

3. So our priors are exchangeable with respect to green.

The first premise says that we are randomly sampling. This means that each emerald has the same probability of being observed at any position in our sequence of observations. I understand random sampling in terms of subjective probability rather than objective chance (Eagle 2005). In a moment, I introduce a simple formal model. In this model, we can formalize random sampling as the

[6]Formally, the random variables $X_1, ..., X_n$ are measurable functions $X_i : \Omega \to \mathbb{R}$ with $1 \le i \le n$ on some probability space $\langle \Omega, \Sigma, \mathbb{P} \rangle$ where Ω is a non-empty set of states, Σ is a σ-algebra on Ω and $\mathbb{P} : \Sigma \to \mathbb{R}$ is a probability function modeling our agent's credences. I often abbreviate $\mathbb{P}(X = x \cap Y = y)$ as $\mathbb{P}(X = x, Y = y)$.

claim that we are equally likely to observe each permutation of the emeralds.[7]

The second premise says that whether emeralds are green is probabilistically independent of when we will observe them. Intuitively, the idea is that learning whether emeralds are green does not give us any information about when we will observe them. Plausibly, green is observation-independent in this sense if we are randomly sampling. If someone tells you how many green and blue emeralds you will observe, you cannot infer anything about the order in which you will observe the emeralds. Note that the observation-independence of green is only plausible under the assumption of random sampling. It is not true in general that whether emeralds are green is probabilistically independent of when they are observed. For example, you might think that green emeralds are closer to the surface and so more likely to be observed early.

An example of a property which is not observation-independent is 'being observed first'. If you learn that an emerald has this property, you can infer that it will be observed first even if you are randomly sampling. Is grue observation-independent? Short answer: No. However, this point deserves more careful consideration so I will discuss it in detail below.

Given that our prior satisfies both random sampling and the observation-independence of green, it must be exchangeable with respect to green. The argument generalizes to any observation-independent property:

> **The Random Principle**: If we are randomly sampling and F is observation-independent, then our prior is exchangeable with respect to F.

Here is a simple model to show how the Random Principle can be formalized. Suppose there are n distinct objects $\mathcal{E} = \{e_1, ..., e_n\}$ and we will exhaustively sample from this population without replacement.[8] Think of states as fixing

[7] Moreland (1976) also discusses random sampling and the New Riddle but uses a different notion of random sampling. Godfrey-Smith (2011) discusses random sampling and the New Riddle in the context of classical statistics. Fitelson and Osherson (2015) and Johannesson (2023) argue that the New Riddle and related puzzles pose problems for classical statistics.

[8] We could adopt this model to allow sampling with replacement and cases where we only partially observe the population. However, I will focus on the model as described here for the sake of tractability.

both which objects are observed when and which objects are F. For example, if $\mathcal{E} = \{e_1, e_2\}$, a state could look as follows: first e_1 is observed, then e_2 is observed, e_1 is F, e_2 is not F.

We can formalize this set-up with the following state space. Let $F = \mathcal{P}(\mathcal{E})$ be all the ways to assign an extension to property F. In our intended interpretation, this corresponds to which emeralds are green. Let O be the set of all sequences of length n of objects in $\mathcal{E}$ without repetition.[9] In our intended interpretation, this corresponds to which order the emeralds are observed in.[10] Then, define our state space $\Omega = F \times O$ and our algebra $\Sigma = \mathcal{P}(\Omega)$. States are pairs $\langle f, o \rangle$, where f specifies a particular extension of F and o specifies a particular order our emeralds will be observed in. We define two random variables $\mathcal{O} : \Omega \to O$ and $\mathcal{F} : \Omega \to F$, where $\mathcal{F}$ is the projection function which returns the first coordinate of any state $\langle f, o \rangle$ and $\mathcal{O}$ is the projection function which returns the second coordinate of any state $\langle f, o \rangle$. Intuitively, $\mathcal{F}$ and $\mathcal{O}$ tell us what the values of F and O at each state are.

Random sampling means that each object in $\mathcal{E}$ has the same probability of being observed at any position, so we are equally likely to observe each sequence of objects:

Definition 2. *We are randomly sampling from $\mathcal{E}$ iff all values of $\mathcal{O}$ are equally likely.*

If F is *observation-independent*, learning the answer to the question 'which objects are F?' does not give us any new information about the question 'when are the objects observed?' and vice versa:

Definition 3. *F is observation-independent iff $\mathcal{F}$ is independent of $\mathcal{O}$, written $\mathcal{F} \perp\!\!\!\perp \mathcal{O}$. This means that for each f in the range of $\mathcal{F}$ and o in the range of $\mathcal{O}$, $\mathbb{P}(\mathcal{F} = f, \mathcal{O} = o) = \mathbb{P}(\mathcal{F} = f)\mathbb{P}(\mathcal{O} = o)$ (Pitman 1993, p. 151).*

[9] Formally, $O = \{\langle e_{\pi(1)}, ..., e_{\pi(n)} \rangle : \{e_1, ..., e_n\} = \mathcal{E}$ and π is a permutation on $\{1, ..., n\}\}$.

[10] As an anonymous referee points out, objects can be observed in many ways. When I say that e is observed, this is shorthand for 'e is observed with respect to being F', which is compatible with e being F and e not being F.

You might think there is a simpler way to define observation-independence. F is *weakly observation-independent* iff $\mathbb{P}(\text{e is F} \mid \text{e is observed in position } j) = \mathbb{P}(\text{e is F})$ for all $e \in \mathcal{E}$ and $1 \leq j \leq n$. If F is observation-independent, then F is also weakly observation-independent. However, it turns out that weak observation-independence is not enough to guarantee exchangeability. An example to show this is in the appendix.

Note that as I understand it, random sampling and observation-independence are conditions on our prior. They are not worldly facts which we may or may not believe to obtain. The Random Principle is given by the following theorem, which I prove in the appendix:

Theorem 1. *Consider the sequence of random variables $X_1, ..., X_n$ where $X_i = 1$ iff the i-th object from $\mathcal{E}$ we observe is F and $X_i = 0$ otherwise. If we are randomly sampling from $\mathcal{E}$ and F is observation-independent, then the sequence $X_1, ..., X_n$ is exchangeable.*[11]

Exchangeability is a familiar concept in Bayesian statistics.[12] However, subjective Bayesians regard exchangeability as a subjective judgment. In contrast, I have explained an objective Bayesian perspective on exchangeability, where we derive exchangeability from random sampling and observation-independence understood as constraints on priors.

[11] Pitman (1993, p. 239) proves a similar result but without making observation-independence explicit. Ericson (1969, p. 197) notes that "[e]xchangeability thus expresses the prior knowledge that while the units of the finite population are identifiable by their labels (here the integers $1, 2, ..., N$) there is no information carried by these labels regarding the associated X_i's", where X_i is the "unknown value of some characteristic possessed by the i-th population element". Diaconis (1977) proves that any finite exchangeable sequence can be viewed as arising from random sampling without replacement, which supports a close connection between exchangeability and random sampling.

[12] The notion of exchangeability was introduced by de Finetti (1937) and independently by Johnson (1924). Similar conditions play an important role in Carnap's inductive logic (Carnap 1950). Good (1969, p. 21) points out that "you would not accept the permutability postulate [exchangeability] unless you already had the notion of physical probability and statistical independence at the back of your mind".

4 All-Grue Is (Very) Improbable

Exchangeability implies an asymmetry between the prior probability of all-green and all-grue. The prior probability of all-green is not constrained. In contrast, there is an upper bound on the prior probability of all-grue. As the total number of emeralds we will observe increases, this upper bound becomes vanishingly low.

Suppose we are about to observe n emeralds by random sampling. All-grue says that for some particular $k < n$, the first k emeralds we will observe are green and the remaining emeralds are blue. So all-grue says that we will observe *one particular ordering* of green and blue emeralds: first all the green ones and then all the blue ones. But there are many other orderings of these green and blue emeralds. To be precise, there are $\binom{n}{k}$ such orderings.[13] Exchangeability says that all of these possible orderings of our green and blue emeralds must have the same prior probability. This means that the maximum prior probability of all-grue is $1/\binom{n}{k}$. Moreover, this upper bound is only attained if we are certain that we will observe exactly k green emeralds. Any uncertainty about the number of green emeralds we will observe decreases the prior probability of all-grue.

As we increase the number of observed emeralds, the upper bound on the prior probability of all-grue quickly becomes very low. Suppose, for example, that all-grue says that the first five emeralds will be green and the remaining emeralds will be blue. If we will observe ten emeralds, the upper bound on the prior probability of all-grue is $1/\binom{10}{5} = 1/252$. If we observe twenty emeralds, the upper bound is $1/\binom{20}{5} = 1/15504$. If we observe fifty emeralds, the upper bound is $1/\binom{50}{5} = 1/2118760$. So if we expect to observe more than a small number of emeralds, exchangeability forces us to assign a vanishingly low prior probability to all-grue.

Here is a quick way to see the asymmetry. Suppose you satisfy exchangeability. It is consistent with that to be certain, in the sense of assigning probability

[13] $\binom{n}{k} = \frac{n!}{k!(n-k)!}$, pronounced '$n$ choose k', is the number of ways to select k objects from n objects. In our example, there are n different positions and k green emeralds which could occupy them.

one, to all emeralds being green. But it is not consistent with that to be certain that all emeralds are grue. Suppose we will observe two emeralds so we have a sequence of two observations X_1, X_2. An emerald is *grue* if it is either observed first and green or observed second and blue. It is consistent with exchangeability to have $\mathbb{P}(11) = 1$ so you can assign probability one to all-green. But all-grue says that $X_1 = 1$ and $X_2 = 0$. Exchangeability requires that $\mathbb{P}(10) = \mathbb{P}(01)$, so the probability of all-grue cannot exceed .5.

Here is another example. Suppose you are about to observe four emeralds. You know that all frequencies of green and blue emeralds are possible: it might be that all emeralds are green, all emeralds are blue, one emerald is green and the rest is blue and so on. One natural way to assign prior probabilities is to consider all five possible 'color frequencies' to be equally likely (Bayes 1763; Carnap 1950). Now consider the all-grue hypothesis which says that the first two emeralds are green and the remaining two emeralds are blue. All-grue says that there are two green emeralds and two blue emeralds arranged in one specific ordering. By our indifference assumption, the probability of two green emeralds and two blue emeralds is $\frac{1}{5}$. And exchangeability requires that all six ways of arranging two green and two blue emeralds are equally likely. Therefore, the prior probability of all-grue is $\frac{\frac{1}{5}}{6} = \frac{1}{30}$ while the prior probability of all-green is $\frac{1}{5}$. I am not committed to this being the uniquely right way to assign prior probabilities. Nonetheless, it shows how once we have exchangeability, it is very easy to end up assigning a lower prior probability to all-grue than to all-green.[14]

Exchangeability explains our intuitive sense that it would be an enormous coincidence if all emeralds were grue while it would not be an enormous coincidence if all emeralds were green. All-grue requires the order of our observations and the color of the emeralds we are observing to be correlated in an extremely

[14]Under exchangeability, we can exactly characterize the conditions under which $\mathbb{P}(\text{all-green}) > \mathbb{P}(\text{all-grue})$ where all-grue says that the first $k < n$ emeralds are green and the rest blue. Assume we will observe n green or blue emeralds modeled by an exchangeable sequence $X_1, ..., X_n$. Let N be the random variable counting the total number of green emeralds observed so $\mathbb{P}(\text{all-green}) = \mathbb{P}(N = n)$ and $\mathbb{P}(\text{all-grue}) = \mathbb{P}(N = k)\frac{1}{\binom{n}{k}}$. $\mathbb{P}(\text{all-green}) > \mathbb{P}(\text{all-grue})$ iff $\mathbb{P}(N = n) > \mathbb{P}(N = k)\frac{1}{\binom{n}{k}}$. Since $\binom{n}{k}$ tends to be large, $\mathbb{P}(\text{all-green}) > \mathbb{P}(\text{all-grue})$ unless you assign *much* higher credence to $\mathbb{P}(N = k)$ than to $\mathbb{P}(N = n)$.

specific way.[15] When randomly sampling, this is very unlikely to happen. So we have a principled reason for assigning a low prior probability to all-grue.

I have explained how exchangeability imposes a strong bias against all-grue. And as we have seen above, exchangeability can be derived from plausible constraints on priors: random sampling and the observation-independence of green. So we have a compelling Bayesian response to the New Riddle.

5 The Symmetry Objection

I have argued that random sampling implies *green exchangeability*: exchangeability with respect to green and blue. But perhaps that is just because I have described the whole set-up in green rather than grue terms. Suppose we assume *grue exchangeability* instead, which is exchangeability with respect to grue and bleen. Grue exchangeability yields an apparently symmetrical argument for an upper bound on the prior probability of all-green.

To address this worry, I first make the idea of grue exchangeability precise. Then, I argue that grue exchangeability does *not* follow from random sampling. My argument for green exchangeability does not work for grue exchangeability because grue is not observation-independent when randomly sampling. I close by reflecting on what it means to respond to the New Riddle and argue that while certain skeptical challenges remain, the New Riddle is less paradoxical than it appears.

5.1 Gruesome Exchangeability

I have described our observations as sequence of random variables $X_1, ..., X_n$, where $X_i = 1$ if the i-th emerald we observe is green and $X_i = 0$ if the i-th

[15] White (2005, p. 19) writes: "On the all-grue hypothesis we have two properties, greenness and having been observed, which are co-instantiated by the same subclass of emeralds. This is a striking fact which seems to call for an explanation. We should not expect apparently causally independent properties to be correlated in this way in a large number of instances." In a similar vein, Chihara (1981, pp. 434–5) writes that "given the sheer number of emeralds discovered thus far [...], it is implausible that the green ones would turn out to coincide exactly with the emeralds observed prior to t_0".

emerald we observe is blue. But we could also describe our observations as sequence of random variables $Y_1, ..., Y_n$, where $Y_i = 1$ if the i-th emerald we observe is grue and $Y_i = 0$ if i-th emerald we observe is bleen. An emerald is *grue* if, for some particular $k < n$, the emerald is observed among the first k emeralds and green or it is not observed among the first k emeralds and blue. An emerald is *bleen* if it is observed among the first k emeralds and blue or it is not observed among the first k emeralds and green. Why not say that $Y_1, ..., Y_n$ is exchangeable? If so, your priors are exchangeable with respect to grue.[16]

Suppose we accept grue exchangeability. Now we can argue that the prior probability of all-green is low. All-green says that there are some grue and some bleen emeralds which will be observed in one specific ordering: first all the grue ones and then all the bleen ones. But there are many other orderings in which these grue and bleen emeralds could be observed. By grue exchangeability, all of these orderings are equally likely. Therefore, there is an upper bound on the prior probability of all-green. In contrast, there are no constraints on the prior probability of all-grue.

Do we have any reason to prefer my argument for the improbability of all-grue over this apparently symmetrical argument for the improbability of all-green? To be clear, there is nothing in the axioms of probability which compels us to accept green exchangeability rather than grue exchangeability—or indeed any kind of exchangeability. However, grue exchangeability does not follow from random sampling.

Let us look at a simple example. We will observe two emeralds known to be either green or blue. An emerald is *grue* if it is either observed first and green or observed second and blue. An emerald is *bleen* if it is either observed first and blue or observed second and green. Grue exchangeability implies that the probability that the first emerald is grue and the second emerald is bleen equals the probability that the first emerald is bleen and the second emerald is grue. In more familiar terms, the probability that both emeralds are green

[16]Kutschera (1978), Earman (1992) and Skyrms (1994) also discuss grue exchangeability but are skeptical about whether gruesome priors can be rationally criticized.

equals the probability that both emeralds are blue. So grue exchangeability forces us to assign the same probability to both emeralds being green and both emeralds being blue. But this constraint does not follow from random sampling. Therefore, we should not accept grue exchangeability.

For example, suppose you are randomly sampling two emeralds which are likely blue. The emeralds being likely blue is perfectly consistent with random sampling. In this case, the probability that both emeralds are blue is higher than the probability that both emeralds are green. In other words, the probability that the first emerald is bleen and the second emerald is grue is higher than the probability that the first emerald is grue and the second emerald is bleen, so grue exchangeability fails.

Here is the reasoning spelled out in more detail. Suppose you are randomly sampling two emeralds a and b. You think the emeralds are more likely blue than green so, for example, $\mathbb{P}(BB) = .6$ and $\mathbb{P}(GG) = .4$.[17] This assignment of probabilities is consistent with random sampling. If we describe this example in terms of our original random variables X_1 and X_2, exchangeability is satisfied. However, if we describe it in terms of our gruesome random variables Y_1 and Y_2, exchangeability fails. This can be seen in figure 1 below. The sequence Y_1, Y_2 is not exchangeable since $\mathbb{P}(Y_1 = 1, Y_2 = 0) = .4$ while $\mathbb{P}(Y_1 = 0, Y_2 = 1) = .6$. What makes this example work is that you assign more probability to the emeralds being blue than to them being green. Such a bias is not essential. As I show in the appendix, there are cases where exchangeability fails for grue-like properties even though emeralds are equally likely to be green and blue.

Event	Probability	Values of X_1, X_2	Values of Y_1, Y_2,
GG	.4	11	10
BB	.6	00	01

Figure 1: Gruesome random variables do not preserve exchangeability.

The upshot: grue exchangeability does not follow from random sampling. While we can translate our whole set-up from familiar green terms to Goodma-

[17]In terms of $\mathcal{F}$ which records which emeralds are green and $\mathcal{O}$, GG is the event $\mathcal{F} = \{a, b\} \cap (\mathcal{O} = \langle a, b \rangle \cup \mathcal{O} = \langle b, a \rangle)$ and BB is the event $\mathcal{F} = \varnothing \cap (\mathcal{O} = \langle a, b \rangle \cup \mathcal{O} = \langle b, a \rangle)$.

nian grue terms, this translation does not preserve exchangeability.

5.2 Breaking the Symmetry

At this point, you might be puzzled. Where does my argument for green exchangeability go wrong for grue exchangeability? What breaks the symmetry between green and grue is that when randomly sampling, green is observation-independent while grue is not.

Before I continue, let me set aside a possible misunderstanding. You might think that grue is observation-dependent because it is explicitly defined in terms of observation. In contrast, green is not explicitly defined in terms of observation. But Goodman (1955, p. 79) points out that this asymmetry depends on our choice of linguistic primitives. If we take grue and bleen as primitive, we can define green in terms of observation: an emerald is green if it is grue and already observed or blue and not yet observed. For this reason, you might think that observation-independence cannot break the symmetry between green and grue. When I say that grue is observation-dependent, I do *not* mean that grue is explicitly defined in terms of observation. Rather, I mean that whether objects are grue is not always probabilistically independent of when they are observed. For this probabilistic notion of observation-independence, it is irrelevant whether we start with green or grue primitives. This means that the property 'either grue and already observed or bleen and not yet observed' can be observation-independent in my sense.

Let me explain why grue is observation-dependent in my sense. Suppose you are about to observe two emeralds by random sampling. An angel appears before you and tells you that you will observe one grue and one bleen emeralds, where an emerald is *grue* if observed first and green or observed second and blue and analogously for *bleen*. Might you have any reason to regard one ordering of the grue and bleen emeralds as more likely? For example, might you have reason to think that you will first observe the grue emerald and then the bleen emerald rather than the other way around? Initially, you might think that the answer is 'no'. How could you use the information the angel has given you to predict anything about the ordering of grue and bleen emeralds?

But appearances are mistaken. You might have reason to think that one ordering of the grue and bleen emeralds is more likely. Suppose you think the emeralds are likely blue. Therefore, you think that the grue emerald is likely observed second. This is because the grue emerald is green if observed first and blue if observed second and you think that the emerald is likely blue. Analogously, you think that the bleen emerald is likely observed first.

We can use the numerical example above to make this point. You randomly sample two emeralds a and b. As before, you think the emeralds are likely blue so $\mathbb{P}(BB) = .6$ and $\mathbb{P}(GG) = .4$. Now suppose you learn that emerald a is grue. This event occurs if either a is observed first and green or a is observed second and blue.[18] Conditional on the information that a is grue, a is more likely to be observed second.[19] Analogously, conditional on the information that a is bleen, a is more likely to be observed first. So in this case, grue and bleen are not even weakly observation-independent. In contrast, green and blue are not observation-dependent in the same way. For example, conditional on the information that a is green, a is equally likely to be observed first or second.[20] As I show in the appendix, there are more complicated examples where grue-like properties are weakly observation-independent but still not observation-independent in my sense.

The upshot: whether emeralds are grue or bleen is not, in general, probabilistically independent of when they are observed. If an angel tells us what colors of emeralds we will observe and we are randomly sampling, we cannot predict anything about their ordering. But if an angel tells us what *schmolors*

[18] In terms of $\mathcal{F}$ and $\mathcal{O}$, the event that a is grue is $(\mathcal{F} = \{a,b\} \cap \mathcal{O} = \langle a,b \rangle) \cup (\mathcal{F} = \varnothing \cap \mathcal{O} = \langle b,a \rangle)$.

[19] By the definition of 'a is grue' and conditional probability, $\mathbb{P}(\mathcal{O} = \langle b,a \rangle \mid a \text{ is grue}) = \frac{\mathbb{P}(\mathcal{F}=\varnothing \cap \mathcal{O}=\langle b,a \rangle)}{\mathbb{P}((\mathcal{F}=\{a,b\} \cap \mathcal{O}=\langle a,b \rangle) \cup (\mathcal{F}=\varnothing \cap \mathcal{O}=\langle b,a \rangle))}$. By finite additivity and observation-independence, this is $\frac{\mathbb{P}(\mathcal{F}=\varnothing)\mathbb{P}(\mathcal{O}=\langle b,a \rangle)}{\mathbb{P}(\mathcal{F}=\{a,b\})\mathbb{P}(\mathcal{O}=\langle a,b \rangle)+\mathbb{P}(\mathcal{F}=\varnothing)\mathbb{P}(\mathcal{O}=\langle b,a \rangle)}$. By random sampling $\mathbb{P}(\mathcal{O} = \langle a,b \rangle) = \mathbb{P}(\mathcal{O} = \langle b,a \rangle) = 0.5$ and by assumption $\mathbb{P}(\mathcal{F} = \{a,b\}) = .4$ and $\mathbb{P}(\mathcal{F} = \varnothing) = .6$ so this is $\frac{0.6 \times 0.5}{0.4 \times 0.5 + 0.6 \times 0.5} = 0.6$.

[20] By the definition of conditional probability, random sampling and observation-independence, $\mathbb{P}(\mathcal{O} = \langle b,a \rangle \mid a \text{ is green}) = \frac{\mathbb{P}(\mathcal{O}=\langle b,a \rangle \cap \mathcal{F}=\{a,b\})}{\mathbb{P}(\mathcal{F}=\{a,b\})} = \frac{\mathbb{P}(\mathcal{O}=\langle b,a \rangle)\mathbb{P}(\mathcal{F}=\{a,b\})}{\mathbb{P}(\mathcal{F}=\{a,b\})} = \frac{0.5 \times 0.5}{0.5} = 0.5 = \mathbb{P}(\mathcal{O} = \langle b,a \rangle)$.

of emeralds we will observe, we might be able to predict something about their ordering even if we are randomly sampling.

5.3 Paradox Lost?

There are ways for the grue skeptic to resist my argument. They can reject random sampling. However, this comes at a cost. As I noted earlier, random sampling seems the best case scenario for inductive inference. So if the grue skeptic rejects random sampling, they can no longer claim that the New Riddle arises even in the best case for inductive reasoning. More generally, rejecting random sampling commits the grue skeptic to think that there is some correlation between the label of an object and when we will observe it, which seems implausible in many circumstances. Furthermore, my definition of random sampling is neutral between green and grue and only talks about objects and which order they are observed in. So it seems unmotivated for the grue skeptic to reject random sampling.

The grue skeptic could also hold that when randomly sampling, green is observation-dependent and grue is observation-independent. This also comes at a cost. Even if you have never heard of grue, you surely think that the colors of emeralds give you no information about when you will observe them when randomly sampling. Under the assumption of random sampling, it seems very implausible to think that we can predict when emeralds are observed on the basis of whether they are green or blue. Believing that you can make such predictions almost seems like believing in magic: an inexplicable correlation between whether emeralds are green and when they are observed. It is not very puzzling that we end up with strange consequences from such a strange starting point. We might not be able to force the grue skeptic to accept the observation-independence of green, but if their position requires us to abandon such a plausible belief, we should not feel threatened by this kind of skeptical challenge. This is particularly plausible if, like Goodman, our project is to find general principles which make sense of our inductive practice.[21]

More broadly, the observation-independence of grue does not follow merely from adopting grue linguistic primitives but is an additional substantive commitment. To be a grue *speaker* does not suffice for being a grue *reasoner*. So there is a substantive disagreement between us and the grue skeptic which is not settled by which language we speak or which linguistic primitives we adopt: the question which properties are observation-independent.

The upshot: while there are ways for the grue skeptic to resist my argument, they come with substantial costs. The grue skeptic seems committed to independently implausible views. Moreover, these views are substantive and not merely the result of adopting different linguistic primitives. So the New Riddle is less puzzling than it initially appears. The New Riddle remains a problem for attempts to define confirmation in syntactical terms, but such theories should be rejected anyways. If you adopt a Bayesian perspective and some plausible substantive assumptions, you need not fear the New Riddle.

While there is room for disagreement about whether my story 'solves' the New Riddle, I think it addresses what is most puzzling about the problem. One might feel puzzled on what grounds we could assign different prior probability to all-green and all-grue. There is a possible world where all emeralds are green and a possible world where all emeralds are grue. What makes the paradox posed by the New Riddle so compelling is not the mere fact that grue-friendly priors are consistent but rather that we seem to have no language-independent reason for discriminating between these possible worlds even in the best case for inductive reasoning. The observation-independence of green provides such a reason and is independently plausible.

This sheds light on the broader problem raised by Goodman. The New Riddle raises the question how inductive reasoning works. Intuitively, we think of

[21] This is a more general point. The disquietude we feel when confronted with skeptical challenges does not stem merely from the fact that the skeptic is consistent and we cannot convince them but that the skeptical argument starts from premises we seem unable to reject. A philosophical response to skepticism must show us how we can reject these premises after all (Stroud 1984).

inductive reasoning as a matter of extrapolating observed patterns. Goodman shows that this intuitive picture cannot be right because "regularities are where you find them, and you can find them anywhere" (Goodman 1955, p. 82). This can leave us puzzled about what we do when we engage in inductive reasoning. We need reasons for distinguishing among different patterns found in our observations. On the view I have offered here, some of these reasons can be found in substantive views about which properties are observation-independent.

If we are evaluating a proposed solution to the New Riddle, the question is not whether the assumptions we make are themselves immune to all skeptical challenges. In particular, it is fair game to use substantive assumptions to break the symmetry between green and grue.[22] Rather, the challenge is to find a language-independent asymmetry between green and grue which is independently plausible and show how this asymmetry constrains our inductive reasoning. As I have explained, the probabilistic notion of observation-independence gets this job done.

Open questions remain. What grounds our belief that some properties are observation-independent while others are not? While this is a deep question, I do not have much to say about it here. Perhaps a more complete account of our inductive reasoning will shed light on why we accept some independence assumptions rather than others. But it might also be a mistake to seek objective grounds for our independence assumptions. Perhaps believing that certain properties are observation-independent is just how we roll. My goal was to show how we can draw on such independence assumptions to explain why the prior probability of all-grue is low.

I do not think that such arguments address all the problems in the vicinity of

[22]The view that we can draw on substantive assumptions to address the New Riddle while not answering all skeptical challenges that can be raised with respect to these assumptions themselves is widely endorsed. Proponents include Goodman (1955), Thomson (1966), Hesse (1969), Jackson (1975), Lewis (1983), Rheinwald (1993), Godfrey-Smith (2002), Sider (2011), Warren (2023), and Zinke (forthcoming). Jackson (1975, p. 131) writes that "[t]his may well raise fundamental problems at the level of justification, in the context of the 'old problem of induction,' but this has not been our concern here." Furthermore, as emphasized by Scheffler (1958) and Stroud (2011, pp. 29–31), the New Riddle poses a problem for describing how we form expectations about the unobserved even if we set aside questions of justification.

the New Riddle. At its most general, the New Riddle is a problem of underdetermination: different theories predict the same evidence. In a Bayesian setting, which theory is supported best depends on your priors.[23] I do not claim that exchangeability and similar independence assumptions completely solve this problem. Inductive reasoning is hard and there are no general-purpose recipes to be had. The best we can hope for is to constrain our inductive inferences in particular situations.

6 Conclusion

To respond to the New Riddle, we must find an asymmetry between green and grue and explain how this asymmetry constrains our inductive reasoning. In a Bayesian framework, this means explaining how the asymmetry constrains our assignment of prior probabilities. The asymmetry should be language independent in the sense of not depending on our choice of linguistic primitives.

My proposal is that green is observation-independent while grue is not, spelled out as probabilistic independence of whether emeralds are green and when they are observed. I have explained how this constrains our priors via the Random Principle. The asymmetry is language independent. So the New Riddle is less paradoxical than it appears. We have gained a better understanding of how to resist the grue skeptic and a more general lesson: independence assumptions are central to our inductive reasoning. This helps to explain how inductive reasoning is a way to gain knowledge of an objective world which is the way it is independently of our observations.

[23]Earman (1992, ch. 6) discusses the problem of the priors. There are general principles for determining prior probabilities such as the principle of indifference or the idea that we should assign higher priors to simpler hypotheses. But these principles are often affected by language dependence and so won't help us to solve the New Riddle without additional reasons for why some languages are better than others (Seidenfeld 1986; Sterkenburg 2016; Piantadosi 2018; Neth 2023). Titelbaum (2010) argues that such problems generalize to any account of confirmation.

Appendix: Proving the Random Principle

Theorem 1. *Consider the sequence of random variables $X_1, ..., X_n$ where $X_i = 1$ iff the i-th object from $\mathcal{E}$ we observe is F and $X_i = 0$ otherwise. If we are randomly sampling from $\mathcal{E}$ and F is observation-independent, then the sequence $X_1, ..., X_n$ is exchangeable.*

Proof. We want to show $\mathbb{P}(X_1 = x_1, ..., X_n = x_n) = \mathbb{P}(X_1 = x_{\pi(1)}, ..., X_n = x_{\pi(n)})$ for any permutation π on $\{1, ..., n\}$. Pick an arbitrary permutation π. We first show that for any f in the range of $\mathcal{F}$ with $\mathbb{P}(\mathcal{F} = f) > 0$, $\mathbb{P}(X_1 = x_1, ..., X_n = x_n \mid \mathcal{F} = f) = \mathbb{P}(X_1 = x_{\pi(1)}, ..., X_n = x_{\pi(n)} \mid \mathcal{F} = f)$. The theorem follows by the law of total probability since the possible values of $\mathcal{F}$ form a partition.[24]

Pick an arbitrary value f of $\mathcal{F}$ with $\mathbb{P}(\mathcal{F} = f) > 0$. Recall that f is a subset of $\mathcal{E}$, namely the set of objects which are F. Now let us analyze $\mathbb{P}(X_1 = x_1, ..., X_n = x_n \mid \mathcal{F} = f)$. The binary sequence $\mathbf{x} = \langle x_1, ..., x_n \rangle$ of values of $X_1, ..., X_n$ contains k ones and $n - k$ zeros with $k \leq n$. This means that k of the objects are F and they are observed in one of the right orderings to produce $\mathbf{x}$.

In terms of our random variables $\mathcal{F}$ and $\mathcal{O}$, we can write the event $X_1 = x_1, ..., X_n = x_n$ as follows. First, k objects are F, so $\mathcal{F} = f$ for some f with $|f| = k$. If $|f| \neq k$, then $\mathbb{P}(X_1 = x_1, ..., X_n = x_n \mid \mathcal{F} = f) = 0 = \mathbb{P}(X_1 = x_{\pi(1)}, ..., X_n = x_{\pi(n)} \mid \mathcal{F} = f)$ and we are done. So assume $|f| = k$.

Second, the objects are observed in one of the right orderings to produce $\mathbf{x}$. Suppose $\mathcal{E} = \{e_1, ..., e_n\}$. We write the value of $\mathcal{O}$ as a sequence $\mathbf{e} = \langle e_1, ..., e_n \rangle$. Say that such a sequence $\mathbf{e}$ agrees with $\mathbf{x}$ relative to $\mathcal{F} = f$ iff for all $1 \leq i \leq n$, $x_i = 1$ iff the i-th object in $\mathbf{e}$ is an element of f and $x_i = 0$ otherwise. The number of sequences which agree with $\mathbf{x}$ is $k!(n-k)!$ since we can permute the objects which are F among themselves and the objects which are not F among themselves without changing the pattern of Fs and not-Fs. (Note that by conditionalizing on $\mathcal{F} = f$ we are holding fixed which objects are F.) Write

[24] As a reminder, the law of total probability says that for any partition $\mathcal{E}$ of our state space Ω and any event A, $\mathbb{P}(A) = \sum_{E \in \mathcal{E}, p(E) > 0} \mathbb{P}(A \mid E)\mathbb{P}(E)$. We show that $\mathbb{P}(A \mid E) = \mathbb{P}(B \mid E)$ for all $E \in \mathcal{E}$ with $\mathbb{P}(E) > 0$. It follows that $\mathbb{P}(A) = \mathbb{P}(B)$

these permutations as $\sigma_1, ..., \sigma_m$ where $m = k!(n-k)!$. For any permutation σ, write $\sigma[\mathbf{e}]$ to abbreviate $\langle e_{\sigma(1)}, ..., e_{\sigma(n)} \rangle$. So we have, by the definition of random variables $X_1, ..., X_n$,

$$\mathbb{P}(X_1 = x_1, ..., X_n = x_n \mid \mathcal{F} = f) = \mathbb{P}(\bigcup_{i=1}^{m} \mathcal{O} = \sigma_i[\mathbf{e}] \mid \mathcal{F} = f) \tag{1}$$

for permutations $\sigma_1, ..., \sigma_m$ chosen as above. By the definition of conditional probability and $\mathbb{P}(\mathcal{F} = f) > 0$,

$$\mathbb{P}(\bigcup_{i=1}^{m} \mathcal{O} = \sigma_i[\mathbf{e}] \mid \mathcal{F} = f) = \frac{\mathbb{P}((\bigcup_{i=1}^{m} \mathcal{O} = \sigma_i[\mathbf{e}]) \cap \mathcal{F} = f)}{\mathbb{P}(\mathcal{F} = f)}.$$

By distributivity and finite additivity, since the events $\mathcal{O} = \sigma_1[\mathbf{e}], ..., \mathcal{O} = \sigma_m[\mathbf{e}]$ are disjoint,

$$\mathbb{P}((\bigcup_{i=1}^{m} \mathcal{O} = \sigma_i[\mathbf{e}]) \cap \mathcal{F} = f) = \sum_{i=1}^{m} \mathbb{P}(\mathcal{O} = \sigma_i[\mathbf{e}] \cap \mathcal{F} = f). \tag{2}$$

By observation-independence $\mathcal{F}$ and $\mathcal{O}$ are independent, so

$$\sum_{i=1}^{m} \mathbb{P}(\mathcal{O} = \sigma_i[\mathbf{e}] \cap \mathcal{F} = f) = \sum_{i=1}^{m} \mathbb{P}(\mathcal{O} = \sigma_i[\mathbf{e}])\mathbb{P}(\mathcal{F} = f). \tag{3}$$

Now we apply the permutation π we picked in the beginning to all sequences of objects $\sigma_1[\mathbf{e}], ..., \sigma_m[\mathbf{e}]$. By random sampling, for any permutation π and sequence $\mathbf{e}$ in the range of $\mathcal{O}$, $\mathbb{P}(\mathcal{O} = \mathbf{e}) = \mathbb{P}(\mathcal{O} = \pi[\mathbf{e}])$, so

$$\sum_{i=1}^{m} \mathbb{P}(\mathcal{O} = \sigma_i[\mathbf{e}])\mathbb{P}(\mathcal{F} = f) = \sum_{i=1}^{m} \mathbb{P}(\mathcal{O} = \pi[\sigma_i[\mathbf{e}]])\mathbb{P}(\mathcal{F} = f). \tag{4}$$

Note that $\pi[\sigma[\mathbf{e}]] = \langle e_{\pi(\sigma(1))}, ..., e_{\pi(\sigma(n))} \rangle$. Reasoning like in our previous steps,

$$\sum_{i=1}^{m} \mathbb{P}(\mathcal{O} = \pi[\sigma_i[\mathbf{e}]])\mathbb{P}(\mathcal{F} = f) = \mathbb{P}((\bigcup_{i=1}^{m} \mathcal{O} = \pi[\sigma_i[\mathbf{e}]]) \cap \mathcal{F} = f). \tag{5}$$

From (2), (3), (4) and (5) we have

$$\mathbb{P}((\bigcup_{i=1}^{m} \mathcal{O} = \sigma_i[\mathbf{e}]) \cap \mathcal{F} = f) = \mathbb{P}((\bigcup_{i=1}^{m} \mathcal{O} = \pi[\sigma_i[\mathbf{e}]]) \cap \mathcal{F} = f). \qquad (6)$$

By the definition of conditional probability and $\mathbb{P}(\mathcal{F} = f) > 0$, (6) entails

$$\mathbb{P}(\bigcup_{i=1}^{m} \mathcal{O} = \sigma_i[\mathbf{e}] \mid \mathcal{F} = f) = \mathbb{P}(\bigcup_{i=1}^{m} \mathcal{O} = \pi[\sigma_i[\mathbf{e}]] \mid \mathcal{F} = f). \qquad (7)$$

We defined $\sigma_1[\mathbf{e}], ..., \sigma_m[\mathbf{e}]$ to be all the sequences which agree with $\mathbf{x}$. So these sequences all have the same pattern of Fs and not-Fs. Then we applied permutation π uniformly to all of these sequences. The resulting sequences $\pi[\sigma_1[\mathbf{e}]], ..., \pi[\sigma_m[\mathbf{e}]]$ also all have the same pattern of Fs and not-Fs but possibly disagree with $\mathbf{x}$. But since we use the same permutation on both sides and hold fixed which objects are F, the sequences $\pi[\sigma_1[\mathbf{e}]], ..., \pi[\sigma_m[\mathbf{e}]]$ are exactly those which agree with $\pi[\mathbf{x}] = \langle x_{\pi(1)}, ..., x_{\pi(n)} \rangle$, so

$$\mathbb{P}(\bigcup_{i=1}^{m} \mathcal{O} = \pi[\sigma_1[\mathbf{e}]] \mid \mathcal{F} = f) = \mathbb{P}(X_1 = x_{\pi(1)}, ..., X_n = x_{\pi(n)} \mid \mathcal{F} = f). \qquad (8)$$

From (1), (7) and (8), we have

$$\mathbb{P}(X_1 = x_1, ..., X_n = x_n \mid \mathcal{F} = f) = \mathbb{P}(X_1 = x_{\pi(1)}, ..., X_n = x_{\pi(n)} \mid \mathcal{F} = f).$$

Since π was an arbitrary permutation, this completes the proof. $\square$

Here is an example to show that weak observation-independence and random sampling do not imply exchangeability. Let $\mathcal{E} = \{a, b, c, d\}$. You are certain that a and b are green and that c and d are blue. Assume you are randomly sampling without replacement so $\mathbb{P}(GGBB) = \mathbb{P}(GBBG) = \mathbb{P}(GBGB) = \mathbb{P}(BBGG) = \mathbb{P}(BGGB) = \mathbb{P}(BGBG) = \frac{1}{6}$. Say that x is F iff x is either observed first and green or x is observed later and blue. Consider the sequence of random variables Y_1, Y_2, Y_3, Y_4 where for $1 \leq i \leq 4$, $Y_i = 1$ iff the i-th observed object is F and $Y_i = 0$ otherwise. We can describe our events in terms of these random variables as shown in figure 2.

Event	Values of Y_1, Y_2, Y_3, Y_4
GGBB	1011
GBBG	1110
GBGB	1101
BBGG	0100
BGGB	0001
BGBG	0010

Figure 2: Redescribing events.

The sequence of random variables Y_1, Y_2, Y_3, Y_4 has the joint distribution $\mathbb{P}(1011) = \mathbb{P}(1110) = \mathbb{P}(1101) = \mathbb{P}(0100) = \mathbb{P}(0001) = \mathbb{P}(0010) = \frac{1}{6}$. The sequence is not exchangeable as $\mathbb{P}(0001) > 0$ but $\mathbb{P}(1000) = 0$.

But F is weakly observation-independent, which means that $\mathbb{P}(x \text{ is } F \mid x \text{ is observed in position } i) = \mathbb{P}(x \text{ is } F)$ for all $x \in \mathcal{E}$ and $1 \le i \le 4$. This is because Y_1, Y_2, Y_3, Y_4 all have the same distribution.[25] This can be seen from the table above since all sequences of values are equally likely, so $\mathbb{P}(Y_1 = 1) = \mathbb{P}(Y_2 = 1) = \mathbb{P}(Y_3 = 1) = \mathbb{P}(Y_4 = 1) = \frac{1}{2}$. So the information that an object is observed in a certain position is independent of whether the object is F. However, my stronger notion of observation-independence is not satisfied. For example, if you learn that there is exactly one F, you know that it will not be observed first.

References

Bayes, Thomas (1763). "An Essay Towards Solving a Problem in the Doctrine of Chances." In: *Philosophical Transactions of the Royal Society of London* 53. With an introduction by R. Price, pp. 370–418. DOI: 10.1098/rstl.1763.0053.

Carnap, Rudolf (1947). "On the Application of Inductive Logic". In: *Philosophy and Phenomenological Research* 8.1, pp. 133–148. DOI: 10.2307/2102920.

— (1950). *Logical Foundations of Probability.* University of Chicago Press.

[25] This example also illustrates that exchangeability of $X_1, ..., X_n$ is stronger than $X_1, ..., X_n$ being identically distributed.

Chihara, Charles (1981). “Quine and the Confirmational Paradoxes”. In: *Midwest Studies in Philosophy* 6.1, pp. 425–452. DOI: 10.1111/j.1475-4975.1981.tb00449.x.

de Finetti, Bruno (1937). “La Prévision: Ses Lois Logiques, Ses Sources Subjectives”. In: *Annales de l'Institut Henri Poincaré* 17, pp. 1–68.

Diaconis, Persi (1977). “Finite Forms of de Finetti's Theorem on Exchangeability”. In: *Synthese* 36, pp. 271–281. DOI: 10.1007/bf00486116.

Eagle, Antony (2005). “Randomness Is Unpredictability”. In: *British Journal for the Philosophy of Science* 56.4, pp. 749–790. DOI: 10.1093/bjps/axi138.

Earman, John (1992). *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*. Bradford.

Elgin, Catherine Z. (1997). *Nelson Goodman's New Riddle of Induction*. Vol. 2. The Philosophy of Nelson Goodman. Taylor & Francis.

Ericson, William A (1969). “Subjective Bayesian Models in Sampling Finite Populations”. In: *Journal of the Royal Statistical Society: Series B (Methodological)* 31.2, pp. 195–224. DOI: 10.1111/j.2517-6161.1969.tb00782.x.

Fitelson, Branden (2008). “Goodman's ‘New Riddle’”. In: *Journal of Philosophical Logic* 37.6, pp. 613–643. DOI: 10.1007/s10992-008-9083-5.

Fitelson, Branden and Daniel Osherson (2015). “Remarks on ‘Random Sequences’”. In: *Australasian Journal of Logic* 12.1. DOI: 10.26686/ajl.v12i1.2134.

Godfrey-Smith, Peter (2002). “Goodman's Problem and Scientific Methodology”. In: *Journal of Philosophy* 100.11, pp. 573–590.

— (2011). “Induction, Samples, and Kinds”. In: *Carving Nature at its Joints: Natural Kinds in Metaphysics and Science*. Ed. by Michael O'Rourke, Joseph Keim Campbell, and Matthew H. Slater. MIT Press, pp. 33–52.

Good, Irving John (1969). “Discussion of Bruno de Finetti's Paper ‘Initial Probabilities: A Prerequisite for any Valid Induction’”. In: *Synthese* 20.1, pp. 17–24. DOI: 10.1007/bf00567233.

Goodman, Nelson (1946). "A Query on Confirmation". In: *Journal of Philosophy* 43.14, pp. 383–385. DOI: 10.2307/2020332.

— (1955). *Fact, Fiction, and Forecast.* Harvard University Press.

Hacking, Ian (1965). *Logic of Statistical Inference.* Cambridge University Press.

Hempel, Carl Gustav (1943). "A Purely Syntactical Definition of Confirmation". In: *Journal of Symbolic Logic* 8.4, pp. 122–143. DOI: 10.2307/2271053.

Hesse, Mary (1969). "Ramifications of 'Grue'". In: *British Journal for the Philosophy of Science* 20.1, pp. 13–25. DOI: 10.1093/bjps/20.1.13.

Hooker, C. A. (1968). "Goodman, 'Grue' and Hempel". In: *Philosophy of Science* 35.3, pp. 232–247. DOI: 10.1086/288211.

Horwich, Paul (1982). *Probability and Evidence.* Cambridge University Press.

Jackson, Frank (1975). "Grue". In: *Journal of Philosophy* 72.5, pp. 113–131. DOI: 10.2307/2024749.

Johannesson, Eric (2023). "The Statistical Riddle of Induction". In: *Australasian Journal of Philosophy* 101.2, pp. 313–326. DOI: 10.1080/00048402.2021.2013909.

Johnson, William Ernest (1924). *Logic, Part III: The Logical Foundations of Science.* Cambridge University Press.

Kutschera, Franz von (1978). "Goodman on Induction". In: *Erkenntnis* 12.2, pp. 189–207. DOI: 10.1007/bf00172226.

Lewis, David (1983). "New Work for a Theory of Universals". In: *Australasian Journal of Philosophy* 61.4, pp. 343–377. DOI: 10.1017/cbo9780511625343.002.

Moreland, John (1976). "On Projecting Grue". In: *Philosophy of Science* 43.3, pp. 363–377. DOI: 10.1086/288693.

Moss, Sarah (2018). *Probabilistic Knowledge.* Oxford University Press.

Neth, Sven (2023). "A Dilemma for Solomonoff Prediction". In: *Philosophy of Science* 90.2, pp. 288–306. DOI: 10.1017/psa.2022.72.

— (ms). "Goodman's New Riddle of Induction Explained in Words of One Syllable". URL: https://philpapers.org/rec/NETGNR.

Norton, John D. (2006). "How the Formal Equivalence of Grue and Green Defeats What is New in the New Riddle of Induction". In: *Synthese* 150.2, pp. 185–207. DOI: 10.1007/s11229-004-6261-z.

Piantadosi, Steven T. (2018). "One Parameter Is Always Enough". In: *AIP Advances* 8.9, p. 095118. DOI: 10.1063/1.5031956.

Pitman, Jim (1993). *Probability.* Springer.

Quine, W. V. (1970). "Natural Kinds". In: *Essays in Honor of Carl G. Hempel.* Ed. by Nicholas Rescher. Dordrecht: D. Reidel, pp. 41–56.

Rheinwald, Rosemarie (1993). "An Epistemic Solution to Goodman's New Riddle of Induction". In: *Synthese* 95.1, pp. 55–76. DOI: 10.1007/bf01064667.

Rosenkrantz, R. D. (1982). "Does the Philosophy of Induction Rest on a Mistake?" In: *Journal of Philosophy* 79.2, pp. 78–97. DOI: 10.2307/2026449.

Scheffler, Israel (1958). "Inductive Inference: A New Approach". In: *Science* 127.3291, pp. 177–181. DOI: 10.1126/science.127.3291.177.

Seidenfeld, Teddy (1986). "Entropy and Uncertainty". In: *Philosophy of Science* 53.4, pp. 467–491. DOI: 10.1086/289336.

Sider, Theodore (2011). *Writing the Book of the World.* Oxford University Press.

Skyrms, Brian (1994). "Bayesian Projectibility". In: *Grue! The New Riddle of Induction.* Ed. by Douglas Stalker. Open Court, pp. 241–62.

Sober, Elliott (1994). "No Model, No Inference: A Bayesian Primer on the Grue Problem". In: *Grue! The New Riddle of Induction.* Ed. by Douglas Stalker. Open Court, pp. 225–240.

Sprenger, Jan and Stephan Hartmann (2019). *Bayesian Philosophy of Science.* Oxford University Press.

Stalker, Douglas Frank (1994). *Grue! The New Riddle of Induction.* Open Court.

Sterkenburg, Tom F. (2016). "Solomonoff Prediction and Occam's Razor". In: *Philosophy of Science* 83.4, pp. 459–479. DOI: https://doi.org/10.1086/687257.

Stroud, Barry (1984). *The Significance of Philosophical Scepticism.* Oxford University Press.

Stroud, Barry (2011). *Engagement and Metaphysical Dissatisfaction: Modality and Value*. Oxford University Press.

Thomson, Judith Jarvis (1966). “Grue”. In: *Journal of Philosophy* 63.11, pp. 289–309. DOI: 10.2307/2024289.

Titelbaum, Michael G. (2010). “Not Enough There There: Evidence, Reasons, and Language Independence”. In: *Philosophical Perspectives* 24.1, pp. 477–528. DOI: 10.1111/j.1520-8583.2010.00201.x.

— (2022). *Fundamentals of Bayesian Epistemology 2: Arguments, Challenges, Alternatives*. Oxford University Press.

Urbach, Peter and Colin Howson (1993). *Scientific Reasoning: The Bayesian Approach*. Open Court.

Warren, Jared (2023). “The Independence Solution to Grue”. In: *Philosophical Studies*, pp. 1–22. DOI: 10.1007/s11098-022-01914-2.

White, Roger (2005). “Explanation as a Guide to Induction”. In: *Philosophers’ Imprint* 5, pp. 1–29. URL: http://hdl.handle.net/2027/spo.3521354.0005.002.

Zinke, Alexandra (forthcoming). “Preservation of Independence and Goodman’s Riddle”. In: *Philosophical Quarterly*. DOI: 10.1093/pq/pqae104.