On The Variable Choice and Multiple Partitions Problems

Adam Edwards Jonathan Livengood

Word Count: 2879

Abstract

We connect the problem of variable choice to the so-called Bertrand Paradox. We begin with a neglected argument by C.S. Peirce against conceptualism, or what we now call "objective Bayesianism." Peirce skewers the conceptualist on a dilemma. Either the conceptualist will not be able to learn from her evidence or she will have to endorse a contradiction. In the argument for the second horn of his dilemma, Peirce produces an instance of Bertrand's Paradox. We consider a version of the Paradox from White (2010) and his argument that the Principle of Indifference is unnecessary to generate Bertrand-style problems. We generalize these arguments and show that any non-trivial partition of the event-space has a paradox-generating counterpart. We observe that partitions are essential to variable definitions and argue that the problem of variable choice—previously thought only a problem for interventionists—is in fact a problem for everyone.

1 Introduction

Anyone who wants to build a model of the world has to decide how to carve things up. In causal modeling, one carves things up by selecting a set of variables and determining the structural relations among them. Hence, causal modelers face the problem of how to choose a set of variables relevant for understanding whatever system they happen to be investigating. Woodward (2016) challenges causal modelers to provide criteria for solving this *problem of variable choice* and specifying how the set of variables ought to be chosen.

In this paper we connect the problem of variable choice to the so-called Bertrand Paradox and to the problem of induction. We begin in Section 2 by considering an argument offered by C.S. Peirce against what he calls "conceptualist"—and we would now call "objective Bayesian"—accounts of probability and statistical inference. Peirce skewers conceptualists on the horns of a dilemma. Either the conceptualist will not be able to learn from her evidence or she will have to endorse a contradiction. In the argument for the second horn of his dilemma, Peirce produces an example of what is now called Bertrand's Paradox. In Section 3, we consider Roger White's take on the paradox and his argument that the Principle of Indifference (POI) is unnecessary to generate Bertrand-style problems. We generalize the Peirce and White arguments in Section 4, by showing that any non-trivial partition of the event-space has a Bertrand-generating counterpart. Finally, in Section 5, we observe that since a partition defines a random variable, the Peirce-Bertrand paradox raises a challenge for objective variable choice that is similar to (or perhaps just an instance of) the problem of induction.

Before we get to the main parts of the paper, however, we want to make a few notes about the Bertrand Paradox and its target. When we start out investigating a new domain, we know very little, if anything, about it. Such severe ignorance could paralyze us. How can we ever get started? One traditional proposal for systematically moving ahead in the face of overwhelming ignorance is expressed by the following:

Principle of Indifference (POI): If your evidence equally supports each of several propositions, then assign the same probability to each.[1]

The POI continues to be the subject of considerable controversy. One standard argument against the POI is fueled by Bertrand's Paradox, which shows that in some cases there are several mutually incompatible ways to assign probabilities over events, each of which is correct according to a naive application of the POI. One reason to continue thinking about Betrand-like challenges to the POI is that the same basic difficulty arises for causal modelers in the guise of the variable choice problem. In cases of total (or near total) ignorance, we do not know how to carve up the world in a way that supports reliable causal inference. Now on to the details.

2 Peirce's Challenge to the Conceptualists

In his 1883 paper, "A Theory of Probable Inference," C. S. Peirce elaborates on an argument that he gave in Peirce (1878) against the conceptualist approach to probability and statistics. The problem Peirce poses is as follows.

Consider an arbitrary binary event-type, E. On any given trial E either occurs or fails to occur, which we encode with 1 and 0, respectively. We represent sequential observations of occurrences or failures as a string of 1s and 0s. For an observation of n trials $\mathbf{E}_n$ is the event-space of all possible outcomes. For example, if there are three trials, then $\mathbf{E}_3 = \{000, 001, 010, 011, 100, 101, 110, 111\}$.

[1] For alternative definitions and discussions of the POI, see Bayes et al. (1763) (but confront Stigler (1982) for an argument that Bayes does not endorse the principle of indifference), Laplace (1774/1986), Keynes (1921, p. 41-42), Barnard (1958), Jaynes (1973), Norton (2008), Bangu (2009), Novack (2010), Vineberg (2011, p.713), Rinard (2014, p. 110), Pettigrew (2016, p. 1), Smithson (2017, p. 255), and Williamson (2018, pp. 560-563).

How should we assign probabilities to these outcomes in situations of total ignorance?

Peirce considers the principle: "if nothing whatever is known about the frequency of occurrence of an event, then any one frequency is as probable as any other."[2] He identifies two ways of proceeding, which form the horns of a dilemma. On the first horn, the conceptualist assigns equal probabilities to each outcome in the event-space but consequently cannot learn from experience. On the second horn, the conceptualist assigns equal probabilities to the possible frequencies of some occurrence but consequently endorses a logical contradiction. Here is a concrete example.

Suppose we have a coin of unknown bias. We flip it three times and observe that it lands heads each time. If we flip the coin again, what is the probability that we will observe heads again? The conceptualist recommends that we proceed by conditionalization as follows:

$$Pr(1111|111) = \frac{Pr(111|1111)Pr(1111)}{Pr(111)}$$

Since $Pr(111|1111) = 1$, the conceptualist only needs to determine the values of $Pr(111)$ and $Pr(1111)$. But how?

The conceptualist could assign equal probabilities to each of the possible outcomes: what Peirce (following Boole) calls "constitutions of the universe." This is the first horn. For a sequence of length three, there are eight possible outcomes—$000, 001, 010, 100, 011, 101, 110, 111$—and each is assigned probability $\frac{1}{8}$. Hence, $Pr(111) = \frac{1}{8}$. For a sequence of length four, there are sixteen possible outcomes, and each is assigned probability $\frac{1}{16}$. Hence, $Pr(1111) = \frac{1}{16}$.

Updating by conditionalization results in a posterior probability of $\frac{1}{2}$, and we would obtain the same result, $\frac{1}{2}$, if we calculated the probability of 1,000

[2] Peirce (1883, p. 172)

heads given 999 so far. Proceeding in this way, the conceptualist cannot learn from experience: regardless of her evidence, she will say that seeing heads on the next flip has probability $\frac{1}{2}$.[3]

We think it is helpful to read the first horn of Peirce's dilemma as making Hume's inductive skepticism formally precise in conceptualist terms. Recall that according to Hume (1748, 7.2.58):

> There appears not, throughout all nature, any one instance of connexion which is conceivable by us. All events seem entirely loose and separate. One event follows another; but we never can observe any tie between them. They seem conjoined, but never connected.

We take it that assigning equal probability to each distinguishable outcome of an experiment—each constitution of the universe—is formally to make all of the outcomes loose and separate. On the first horn, every outcome resembles every other outcome to the exact same degree. We can't, on this horn of the dilemma, associate any impression with any other impression since there is nothing in our experience or in our reason that genuinely connects them. Thus, when we attempt to use Bayes' theorem to update our beliefs on the basis of our experiences we find that we are perfect inductive skeptics! The first horn is unacceptable. But as we'll see, the second horn is worse.

On the second horn of the dilemma, the conceptualist assigns equal probabilities to particular *frequencies* of outcomes. The difficulty for the conceptualist is determining precisely *which* frequency is the appropriate one. Peirce writes: "It will be seen that different frequencies result some from more and some from

[3] In Chapter 20 of his *Investigations of the Laws of Thought*, Boole argued for the first horn of Peirce's dilemma, explicitly drawing the conclusion that "past experience does not in this case affect future expectation" (1854, 371-372). In his *Treatise on Probability*, Keynes (1921, 50, note 3) incorrectly claimed that Peirce had *endorsed* the idea of assigning equal probabilities to all constitutions of the universe. Carnap (1952, 39-40) cited Keynes' claim without correcting the error and went on to note that the rule leads to unacceptable results. See Fitelson (2006a,b) for more discussion of the problem.

fewer different "constitutions of the universe," so that it is a very different thing to assume that all frequencies are equally probable from what it is to assume that all constitutions of the universe are equally probable."[4] Consider the following two options.

First, we could count the frequency of 1s in a sequence. We represent these frequencies in Figure 1.

$F_{0,3}$	$F_{1,3}$	$F_{2,3}$	$F_{3,3}$
000	001	011	111
	010	101	
	100	110	

$F_{0,4}$	$F_{1,4}$	$F_{2,4}$	$F_{3,4}$	$F_{4,4}$
0000	0001	0011	0111	1111
	0010	0101	1011	
	0100	1001	1101	
	1000	0110	1110	
		1010		
		1100		

Figure 1: Frequency of heads for three trials ($F_{X,3}$) and four trials ($F_{X,4}$).

If we assign equal probabilities to these frequencies $Pr(000) = Pr(001 \vee 010 \vee 100) = Pr(011 \vee 101 \vee 110) = Pr(111) = \frac{1}{4}$ and $Pr(0000) = Pr(0001 \vee 0010 \vee 0100 \vee 1000) = \ldots = Pr(1111) = \frac{1}{5}$ then we obtain the posterior probability $Pr(1111|111) = \frac{Pr(1111)}{Pr(111)} = \frac{4}{5}$. The conceptualist now appears to be learning from experience, which seems like an improvement. But there is a problem. As Peirce points out, there are many possible constitutions of the universe, and we cannot consistently assign equal probabilities to the frequencies of all such constitutions at once. To see this, consider the frequency of *changes* internal to a sequence, such as from a run of 0s to a run of 1s, as shown in Figure 2.

On this way of carving things up, $Pr(0000) = Pr(0101) = \frac{1}{8}$, $Pr(1000) = \frac{1}{24}$, etc. We could stop at this point, observe that $\frac{1}{8} \neq \frac{1}{5}$, and conclude that the two ways of assigning probabilities are inconsistent. But Peirce goes one step further and returns to the inference problem. What will the two ways of carving things up say about the probability that the next flip will come up heads? Dividing up

[4]Peirce (1883, p. 173)

$C_{0,3}$	$C_{1,3}$	$C_{2,3}$
000	001	010
111	100	101
	110	
	011	

$C_{0,4}$	$C_{1,4}$	$C_{2,4}$	$C_{3,4}$
0000	0001	0010	0101
1111	0011	0100	1010
	0111	0110	
	1110	1001	
	1100	1011	
	1000	1101	

Figure 2: Frequency of changes for three trials ($C_{X,3}$) and four trials ($C_{X,4}$).

the event-space by *number of heads* gives us a posterior probability of $\frac{4}{5}$ whereas dividing it up by *number of changes* gives us $\frac{3}{4}$. Applying the POI in this way leads the conceptualist to endorse a contradiction.

3 The Multiple Partitions Problem

What has gone wrong? Peirce and others have thought that cases like this illustrate a problem with the Principle of Indifference.[5] However, White (2010) has a beautiful little argument that the problem has nothing special to do with the Principle of Indifference. White's argument is based on consideration of a slightly different example from van Fraassen (1989).

White (2010) calls his example *Mystery Square*:[6]

> A mystery square is known only to be no more than *two* feet wide. Apart from this constraint, you have no relevant information concerning its dimensions. What is your credence that it is less than *one* foot wide? (p. 164)

White says we have no more reason to suppose the square is less than one

[5] Sometimes called the Principle of Insufficient Reason.

[6] In van Fraassen (1989), the example involves cubes: "Consider a factory that produces cubes with edge lengths of no more than two meters. A cube from the latest batch is selected. What should your credence be that the cube has an edge length of less than one meter? Likewise, what should your credence be regarding the face area or volume of the cube?" See van Fraassen (1989, pp. 302–307) for this and other examples.

foot wide than that it is more than one foot wide, or that its area is less than 1 square foot, or between 1 and 2 square feet, and so on. These options are represented in Figure 3.

Edge Length	**Face Area**
L_1: $0m <$ length $\leq 1m$	A_1: $0m^2 <$ area $\leq 1m^2$
L_2: $1m <$ length $\leq 2m$	A_2: $1m^2 <$ area $\leq 2m^2$
	A_3: $2m^2 <$ area $\leq 3m^2$
	A_4: $3m^2 <$ area $\leq 4m^2$

Figure 3: Two partitions over the set of possible squares.

White defines evidential symmetry as follows. For any two propositions X and Y, say that X is *evidentially symmetric* to Y, denoted $X \approx Y$, for some agent if the agent's evidence offers no more support to X than it does to Y and vice versa. Given his definition, White formalizes the POI as shown in 1.

$$X \approx Y \rightarrow Pr(X) = Pr(Y) \tag{1}$$

Plausibly, each partition in Figure 3 is equally legitimate. However, when we attempt to assign probabilities to these partitions we find that we can't do so in a consistent way. White sets out an argument via contradiction to this effect (Figure 4).

(A1) $L_1 \approx L_2$ (premise)
(A2) $A_1 \approx A_2 \approx A_3 \approx A_4$ (premise)
(A3) $Pr(L_1) = 0.5$ (from POI, A2)
(A4) $Pr(A_1) = 0.25$ (from POI, A3)
(A5) $L_1 = A_1$ (equivalence)
(A6) $\therefore Pr(L_1) = Pr(A_1)$ (*contradiction*)

Figure 4: White's argument that POI leads to a contradiction

So far, we just have a variant on the Peirce-Bertrand paradox. But White goes on to argue that we obtain a worrying result without the POI if we assume two plausible principles regarding evidential symmetries.

Transitivity: If $X \approx Y$ and $Y \approx Z$, then $X \approx Z$.

Equivalence: If two propositions are logically equivalent, then they are evidentially symmetric.

With the principles of Transitivity and Equivalence, White derives something seemingly absurd. The argument is given in Figure 5.

(B1) $L_1 \approx L_2$ (premise)
(B2) $A_1 \approx A_2 \approx A_3 \approx A_4$ (premise)
(B3) $\therefore L_1 \approx A_1$ (by equivalence)
(B4) $\therefore L_2 \approx (A_2 \vee A_3 \vee A_4)$ (by equivalence)
(B5) $\therefore A_2 \approx (A_2 \vee A_3 \vee A_4)$ (by transitivity)

Figure 5: White's Multiple Partitions Problem

The B-argument has the same premises as the A-argument, and it concludes with what White calls an "obviously wrong" result. White's proof does not rely on the principle of indifference, but the premises appear likewise motivated by symmetry reasoning. Each partition of possible outcomes seems "as good as the other."

But why think the result of the argument in Figure 5 is bad? The reasoning seems to go something like this. Consider any two possible outcomes for some event. If one possible outcome is a subtype of the other, then the two outcomes cannot be evidentially symmetric. If they *were*, then we could conclude that the difference between them is irrelevant. However, we cannot conclude that the difference between them is irrelevant given our state of total ignorance.[7]

[7] Whether or not we are in a position to know $X \subset Y$ depends on what we mean by being in a state of "total ignorance" with respect to the question at hand. It appears that White thinks our ignorance does not exclude our being in such a position.

4 Generalization of the Problems

In this section we draw some general conclusions about the Bertrand paradox from what we've seen so far. To do so, we will need some technical vocabulary for understanding set partitions.

Partition: Consider a set $E = \{e_1, \ldots e_n\}$. A *partition* X on E is a collection of subsets of E where[8]

1. If $x \in X$ then $x \neq \emptyset$
2. If $x_i \in X$ and $x_j \in X$ then $x_i = x_j$ or $x_i \cap x_j = \emptyset$
3. $\bigcup_{x \in X} X = E$

Refinement: Consider a set $E = \{e_1, \ldots e_n\}$ and two partitions over E: X and Y. X is a *refinement* of Y, denoted $X \preceq Y$, iff $\forall x \in X \exists y \in Y : x \subseteq y$. Call X a *strict* refinement of Y, denoted $X \prec Y$, if $X \neq Y$.

Partial Refinement: Let $E = \{e_1, \ldots, e_n\}$ be a set with two partitions X and Y. X is a *partial* refinement of Y, denoted $X \precsim Y$, iff $\forall x \in X \exists y \in Y : \bigcup_i^n X_i = \bigcup_j^m Y_j \neq E$

Complete Partition: A partition of a countable set is *complete* if each element of the partition is a singleton set.

While there are many ways in which one partition may be a strict refinement of another, a simple procedure for generating a partition Y that is a strict refinement of another partition X is as follows. Begin with partition X and then split at least one element of X into two disjoint, nonempty subsets. Any new partition created this way is a strict refinement of the original.

If we assume that the event space is finite ($E = \{e_1, \ldots, e_n\}$), there is a fixed number of partitions of E given by the Bell sequence.[9] This sequence lists the

[8]This definition appears in Smith et al. (2014).
[9]OEIS Foundation Inc. (2019)

number of unique partitions B_n for sets containing n elements. For any finite set, there exists a partition and a refinement of that partition that will generate a Bertrand paradox. Here is the proof.

Consider a partition X of E. For any partition of a finite set, either a strict refinement of that partition exists or the partition is complete. So, for any non-complete partition X a strict refinement $Y \prec X$ exists. Since $X \neq Y$, when we assign probabilities uniformly over the elements in X and Y the probabilities will be non-identical for some elements of X and Y.

(C1) $X_1 \approx X_2 \approx \ldots \approx X_n$
(C2) $Y_1 \approx Y_2 \approx \ldots \approx Y_m$
(C3) $\forall y \in Y(y \approx (\bigcup_i^{i+n} X_i))$ (by equivalence)
(C4) $(y_p \approx (\bigcup_i^{i+n} x_i))$ (by instantiation)
(C5) $(y_q \approx (\bigcup_j^{j+m} x_j))$ (by instantiation)
(C6) $\therefore (\bigcup_i^{i+n} x_i) \approx (\bigcup_j^{j+m} x_j)$ (by transitivity)

Figure 6: Generalized Multiple Partitions Problem

White's case involves partitioning an uncountable event-space.[10] In *Mystery Square* White proposes two partitions of the size of the square: one based on length and the other on area. He then assigns probability uniformly over the two partitions L and A since each element of the partition is presumed to be evidentially symmetric with each other element of its respective partition.

Because one partition refines the other—in this case $A \prec L$—assigning probabilities uniformly over these partitions generates a paradox. While a Bertrand paradox is not generated for every possible pair of partitions, for any choice of partition that is not complete there exists another partition such that the pair generates a Bertrand paradox.

On the first horn of Peirce's dilemma, we take the complete partition of the event-space and uniformly distribute probability over every elementary event.

[10]This is also the case in Bertrand's original statement of the paradox.

As we've seen, doing so makes learning from experience impossible. Thus, Peirce's argument has shown that learning is impossible if you assign probability uniformly over the complete partition of the event space.

On the second horn of the dilemma, Peirce challenges the conceptualist to identify which of the other possible partitions over the event-space is the correct one. While he only considers two partitions, this horn shows that we can't rationally choose *any* partition, since for any non-complete partition there exists a paradox-generating refinement.

Furthermore, neither of the partitions Peirce considers in his coin flip example (number of heads and number of changes) are refinements of the other. In fact, because partition C is only a partial refinement of partition F, Peirce's coin flip generates the same results as in White's example for three flips but not four.

(D1) $F_{0,3} \approx F_{1,3} \approx F_{2,3} \approx F_{3,3}$
(D2) $C_{0,3} \approx C_{1,3} \approx C_{2,3}$
(D3) $C_{0,3} \approx (F_{0,3} \vee F_3)$ (by equivalence)
(D4) $(C_{1,3} \vee C_{2,3}) \approx (F_{1,3} \vee F_{2,3})$ (by equivalence)
(D5) $C_{1,3} \approx (C_{1,3} \vee C_{2,3})$ (by transitivity)
(D6) $\therefore (C_{0,3} \vee C_{1,3}) \approx (C_{0,3} \vee C_{1,3} \vee C_{2,3})$ (by symmetry preservation)

Figure 7: White's argument applied to three coin flips

No choice of partition is worry free. A complete partition does not result in contradiction, but it also doesn't support learning. But partitions that are not complete endorse outright contradictions (of the Bertrand variety) or absurdities (of the White variety).

5 The Variable Choice Problem

So far we've seen a generalization of the Bertrand paradox according to which any event space can be partitioned and that partition given a refinement that will generate a paradox. In this section we connect Bertrand's paradox to the problem of variable choice.

Woodward (2016, p. 1048) states the problem as follows: "[Consider] a situation in which we can construct or define new previously unconsidered variables either *de novo* or by transforming or combining or aggregating old variables, and where our goal is to find variables that are best or most perspicuous from the point of view of causal analysis / explanation." The worry, for Woodward, is that while this is a task which all modelers must accomplish there is nothing systematic to be said about how best to perform it.

There are two ways to understand the variable choice problem, depending on what we think a variable is. On one understanding a variable is just a property that relates to some piece of mathematics. The problem of variable choice then becomes the problem of determining which properties we ought to include in our model. Another way of understanding variables is as a particular representation of a property. In that case, the problem of variable choice becomes a problem of *how to represent* properties, rather than which properties to choose in the first place.

Consider the following problem. Suppose we are interested in the relationship between smoking behavior and the appearance of lung cancer in later life. A simple way of proceeding would be to define two variables (corresponding to our two properties) that take a binary value (smoked/has never smoked and has lung cancer/no lung cancer). On the first reading, we ask which other properties might be relevant in this context (such as occupation, etc.). On the second reading, we ask whether binary values are sufficient for our model. Perhaps the ‘smoking’ variable should take an integer value that corresponds to the total cigarettes smoked, or a real value that corresponds to the average number of cigarettes smoked per week, etc. Likewise for the ‘Lung Cancer’ variable. Before we even collect data we will need to know the possible values that these variables can take.

The process of identifying variables necessarily involves partitioning our observations in some way. But as we’ve seen, there is no worry-free way to do so in a state of ignorance. An easy skeptical argument follows. We begin in a state of ignorance. But there is no worry-free way to identify variables in such a state. If so, then we should be suspicious that any model gets the “correct” partition of the world.[11]

6 Conclusion

In this paper we considered some of the history of Bertrand’s paradox and the way variations on the paradox have been used to challenge the POI. We generalized arguments developed by Peirce and White, and we connected these Bertrand-like paradoxes to the problem of variable choice. We then presented a skeptical challenge to modelers (especially but not exclusively causal modelers) who need to carve the world into parts in order to make reliable inferences.

References

Sorin Bangu. On Bertrand’s paradox. *Analysis*, 70(1):30–35, 2009.

GA Barnard. Studies in the history of probability and statistics. *Biometrika*, 45(3/4):293–315, 1958.

[11] One might be tempted to draw an even stronger conclusion: There is no correct way to partition an event space, and by parity of reasoning, there is no correct way to define variables, causal or otherwise.

Thomas Bayes, Richard Price, and John Canton. An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, 53:370–418, 1763.

Branden Fitelson. Logical foundations of evidential support. *Philosophy of Science*, 73(5):500–512, 2006a.

Branden Fitelson. The paradox of confirmation. *Philosophy Compass*, 1(1): 95–113, 2006b.

Edwin T Jaynes. The well-posed problem. *Foundations of Physics*, 3(4):477–492, 1973.

John Maynard Keynes. *A Treatise on Probability*. MacMillan and Co., 1921.

Pierre Simon Laplace. Memoir on the probability of the causes of events. *Statistical Science*, 1(3):364–378, 1774/1986.

John D Norton. Ignorance and indifference. *Philosophy of Science*, 75(1):45–68, 2008.

Greg Novack. A defense of the principle of indifference. *Journal of Philosophical Logic*, 39(6):655–678, 2010.

OEIS Foundation Inc. The On-Line Encyclopedia of Integer Sequences, 2019. URL `https://oeis.org/A000110`. (accessed February 5, 2019).

Charles Sanders Peirce. The probability of induction. *Popular Science Monthly*, 12(45):705–718, 1878.

Charles Sanders Peirce. A theory of probable inference. *The Johns Hopkins Studies in Logic*, pages 126–181, 1883.

Richard G. Pettigrew. Accuracy, risk, and the principle of indifference. *Philosophy and Phenomenological Research*, 92(1):35–59, 2016.

Susanna Rinard. The principle of indifference and imprecise probability. *Thought: A Journal of Philosophy*, 3(2):110–114, 2014.

Douglas Smith, Maurice Eggen, and Richard St. Andre. *A Transition to Advanced Mathematics*. Nelson Education, 2014.

Robert Smithson. The principle of indifference and inductive scepticism. *British Journal for the Philosophy of Science*, 68(1):253–272, 2017.

Stephen M Stigler. Thomas Bayes's Bayesian inference. *Journal of the Royal Statistical Society. Series A (General)*, pages 250–258, 1982.

Bas C. van Fraassen. *Laws and Symmetry*. Oxford University Press, 1989.

Susan Vineberg. Paradoxes of probability. In *Philosophy of Statistics*, pages 713–736. Elsevier, 2011.

Roger White. Evidential symmetry and mushy credence. In T. Szabo Gendler and J. Hawthorne, editors, *Oxford Studies in Epistemology*, volume 3, pages 161–186. Oxford University Press, 2010.

Jon Williamson. Justifying the principle of indifference. *European Journal for Philosophy of Science*, pages 1–28, 2018.

James Woodward. The problem of variable choice. *Synthese*, 193(4):1047–1072, 2016.