Scientific Explanation — Final Set of Background Notes

- Administrative:
  - I've added some readings on *Inductive Logic* to our background.
- Inductive Logic (and "Logical Probability")
  - "Logical Probability" and Carnapian Inductive Logic
  - Carnap's (and Hempel's) Dream about Inductive Logic
  - Why The Dream didn't come true.
  - Some General Remarks on (Formal) Logic and Epistemology
- More historical background (Salmon) leading up to Woodward's book
  - Hempel's I–S Model
  - Salmon's S–R Model

Subjective Theories of Probability II

- It seems clear that there is such a thing as "degree of belief". And, it also seems clear that there are *some* sorts of constraints on such degrees.
- But, should degrees of belief obey the *probability axioms*? If so, *Why*?
- There are arguments to the effect that epistemic (accuracy arguments) and pragmatic (dutch book arguments) dob's should be *probabilities*.
  - I will not be discussing such arguments in the seminar.
- Most commentators who write about scientific explanation seem to think that subjective probabilities (epistemic or pragmatic) have no place in any adequate account of explanation. We'll return to this below (Coffa).
- Moreover, so-called "logical" probabilities (if there be such) also seem out of place here, since explanatory relations do not seem to be "logical".
- However, *if* there is such a thing as *inductive logic* (*and it is grounded in "logical probabilities"*!) *then* a Hempelian might have use for them after all — in order to gauge "inductive strength" of *inductive arguments*.

Whither Logical Probability? 1

- The motivation for "logical interpretations" of probability has its roots in *inductive logic*. Inductive Logic (more later) is meant to be the science of *argument strength* — a quantitative generalization of entailment ($\vDash$).
- Inductive Logic aims to explicate a quantitative measure $\mathfrak{c}(C, P)$ of the "degree to which the premises of an argument P jointly confirm its conclusion C." Keynes, Carnap, and others worked on such theories.
- As we will study more in our confirmation unit, one of the key desiderata of IL is that $\mathfrak{c}$ should quantitatively generalize deductive entailment:

($\mathcal{D}_1$) The relations of deductive entailment and deductive refutation should be captured as limiting (extreme) values of $\mathfrak{c}$ with cases of 'partial entailment' and 'partial refutation' lying somewhere on a $\mathfrak{c}$-continuum (or range) between these extreme values of the confirmation function.

- Exercises: (i) *the conditional probability* [$\Pr(C \mid P)$] satisfies $\mathcal{D}_1$, but *the probability of the corresponding conditional* [$\Pr(P \to C)$] does *not*.

Whither Logical Probability? 2

- Another key historical desideratum for inductive logic is:

($\mathcal{D}_2$) Inductive logic (*i.e.*, the *non*-deductive relations characterized by inductive logic) should be *objective* and *logical*.

- Carnap on this desideratum:

> "Deductive logic may be regarded as the theory of the relation of logical consequence, and inductive logic as the theory of another concept [$\mathfrak{c}$] which is likewise objective and logical, *viz.*, ... degree of confirmation."

- Carnap on the terms "logical" and "objective" as they apply to $\mathfrak{c}$.

> "The principal common characteristic of the statements in both fields is their independence of the contingency of facts. This characteristic justifies the application of the common term 'logic' to both fields."
>
> "That $\mathfrak{c}$ is objective means this: if a certain $\mathfrak{c}$ value holds for a hypothesis (H) with respect to evidence (E), then the value $\mathfrak{c}(H, E)$ is independent of what any person may happen to think about these sentences."

Whither Logical Probability? 3

- While $\Pr(C \mid P)$ clearly satisfies $\mathcal{D}_1$, it is unclear whether $\Pr(C \mid P)$ satisfies $\mathcal{D}_2$. If $\mathfrak{c}(C, P) \stackrel{\text{def}}{=} \Pr(C \mid P)$, then, we seem to need $\Pr(C \mid P)$ *itself* to be logical if $\mathfrak{c}(C, P)$ is to be logical. Hence, the need for "logical probability"!
- Moreover, if we assume the standard definition of $\Pr(C \mid P)$, then we need an unconditional probability function that is itself logical. The basic "Leibnizian" idea, which underlies this "interpretation" of probability:

$$\Pr(C \mid P) = \frac{\Pr(P \,\&\, C)}{\Pr(P)} = \frac{\text{The proportion of logically possible worlds in which } P \,\&\, C \text{ is true}}{\text{The proportion of logically possible worlds in which } P \text{ is true}}$$

- Such unconditional logical probability functions are called logical *measure functions* ($\mathfrak{m}$). Intuitively, these are intended to measure the "proportion of logically possible worlds in which a proposition is true".
- Wittgenstein, Carnap, and others give precise explications of this vague concept of "logical probability". They work within logical *languages* $\mathcal{L}$, and they work with *descriptions* of possible worlds — *sentences* in $\mathcal{L}$.

Wittgensteinian Propositional Logical Probability 1

- In the *Tractatus*, Wittgenstein presents the idea of a truth-table for a propositional language $\mathcal{L}_P$. He proposes a logical measure function $\mathfrak{m}$ on $\mathcal{L}_P$, which assigns *equal probability* to each *state description* of $\mathcal{L}_P$.
- Let $\mathcal{L}_P^n$ be a propositional language with n atomic sentences. The measure function $\mathfrak{m}$ would assign $\mathfrak{m}(s_i) = \frac{1}{2^n}$, for all state descriptions s_i of $\mathcal{L}_P^n$. Here's what $\mathfrak{m}$ looks like over an $\mathcal{L}_P^3$ language (this should look familiar!):

A	B	C	State Descriptions	$\mathfrak{m}(s_i)$
T	T	T	$s_1 = A \,\&\, B \,\&\, C$	$\mathfrak{m}(s_1) = 1/8$
T	T	F	$s_2 = A \,\&\, B \,\&\, {\sim}C$	$\mathfrak{m}(s_2) = 1/8$
T	F	T	$s_3 = A \,\&\, {\sim}B \,\&\, C$	$\mathfrak{m}(s_3) = 1/8$
T	F	F	$s_4 = A \,\&\, {\sim}B \,\&\, {\sim}C$	$\mathfrak{m}(s_4) = 1/8$
F	T	T	$s_5 = {\sim}A \,\&\, B \,\&\, C$	$\mathfrak{m}(s_5) = 1/8$
F	T	F	$s_6 = {\sim}A \,\&\, B \,\&\, {\sim}C$	$\mathfrak{m}(s_6) = 1/8$
F	F	T	$s_7 = {\sim}A \,\&\, {\sim}B \,\&\, C$	$\mathfrak{m}(s_7) = 1/8$
F	F	F	$s_8 = {\sim}A \,\&\, {\sim}B \,\&\, {\sim}C$	$\mathfrak{m}(s_8) = 1/8$

Wittgensteinian Propositional Logical Probability 2

- Here are some important *technical* facts about Wittgensteinian $\mathcal{L}_P^n$-Pr:
  - Every atomic sentence in every $\mathcal{L}_P^n$ probability model has probability $\frac{1}{2}$.
  - Every collection of atomic sentences $\{A_1, \ldots, A_n\}$ in every $\mathcal{L}_P^n$ probability model is mutually probabilistically independent.
  - For every pair of sentences p and q (atomic or compound) in every $\mathcal{L}_P^n$ probability model, p and q are probabilistically dependent *iff* p and q are logically dependent (*i.e.*, *iff* $p \vDash q$ or ${\sim}p \vDash q$ or $q \vDash p$ or ${\sim}q \vDash p$). [This depends on the *equiprobability* of the state descriptions. *Why*?]
  - For all p and q in any $\mathcal{L}_P^n$ model, $\Pr(q \mid p) = 1$ if and only if $p \vDash q$. [This does *not* depend on the *equiprobability* of the S.D.'s. *Why*?]
- The key philosophical question about this kind of approach is:
  - What makes the assignment of *equal* probability to the *state descriptions the* "logical" probability assignment? What *logical* principle is at work here? There are 2 kinds of answers: (i) the *Principle of Indifference* (PI), and (ii) *permutation invariance*. Carnap on (PI)...

Carnap on the Principle of Indifference 1

- The Principle of Indifference (PI) says, roughly, that if an epistemically rational agent's total evidence K does not favor any member of a partition of possible states $\{s_1, \ldots, s_n\}$ over any other member, then that agent's degrees of credence (*assumed* to be probabilities!) should satisfy:

$$\Pr(s_i \mid K) = \Pr(s_j \mid K), \text{ for all } i, j.$$

- The invocation of (PI) in this context should be puzzling for a few reasons:
  - First, (PI) *sounds like* an *epistemic* principle about what an agent's epistemic probabilities should be, under certain circumstances.
  - Should *logical* principles be determined by *epistemic* constraints? This is controversial, even for deductive logic (think inconsistent beliefs):
    - ∗ You shouldn't *believe everything* whenever your set of beliefs happens to be inconsistent (remember The Preface Paradox!).
  - So, even for *de*ductive logic, *epistemic* principles seem to make for odd *logical* constraints. Why should *in*ductive logic be any different?

- Presumably, the (PI) approach goes as follows. Assume that "all we know is logic". Call this background knowledge $K_\top$. Assume that $K_\top$ does not favor any state description over any other. Then, $\mathfrak{m}$ can be thought of as $\Pr(\bullet \mid K_\top)$, and we have $\Pr(s_i \mid K_\top) = \Pr(s_j \mid K_\top)$, for all i, j, as desired.
- Carnap cleverly argues that (PI) is a *logical* (*not* epistemic!) principle:

  > ... the statement of equiprobability to which the (PI) leads is, like all other statements of inductive probability, not a factual but a logical statement. If the knowledge of the observer does not favor any of the possible events, then with respect to this knowledge as evidence they are equiprobable.

- Here's how the argument is supposed to go (as far as I can see):
  - Intuitively, "K does not *favor* s_i over s_j (any $i \neq j$)" ⇒ "K *confirms* s_i to the same degree as K *confirms* s_j (each $i \neq j$)", *i.e.*, $\mathfrak{c}(s_i, K) = \mathfrak{c}(s_j, K)$.
  - If we *assume* that $\mathfrak{c}(x, y) = \Pr(x \mid y)$, then *logic alone* implies that:
    $$\mathfrak{c}(s_i, K) = \mathfrak{c}(s_j, K) \Rightarrow \Pr(s_i \mid K) = \Pr(s_j \mid K), \text{ for } all\ i \text{ and } j.$$
  - ∴ *Logic alone* implies that if K does not favor s_i over s_j (for any $i \neq j$ in a partition of states), then $\Pr(s_i \mid K) = \Pr(s_j \mid K)$, for each i, j. □

Carnap on the Principle of Indifference 2

- Carnap's argument is clever, but ultimately not terribly compelling.
- As we will see next, Carnap's own Predicate-Logical approach to "logical probability" deviates from the application of (PI) to the state descriptions of his logical languages [and, for reasons that look *epistemic*, yet again!].
- Aside from these worries about the "logicality" of (PI) and its success at generating the "right 'logical' probability models", there are other problems with Carnap's argument. Here's a cursory sketch (more later):
  - Carnap *assumes* that the correct (probabilistic) explication of the *confirmation* function is just the conditional probability function, *i.e.*,
    $$\mathfrak{c}(x, y) = \Pr(x \mid y)$$
  - If it were obvious that $\Pr(x \mid y)$ is the best (probabilistic) explication of $\mathfrak{c}(x, y)$, then Carnap's argument might be sound (if not informative!).
  - We'll soon see that $\Pr(x \mid y)$ may *not* be the best explication of $\mathfrak{c}(x, y)$.
- This same set of issues arises in the context of Explanation (below)...

Carnapian Monadic Predicate Logical Probability 1

- Generalizing on Wittgenstein, Carnap defined his logical measure functions over sentences in monadic predicate logical languages $\mathcal{L}_Q^{m,n}$ containing n monadic predicates (F, G, ...) and m constants (a, b, ...).
- To fix ideas, consider the language $\mathcal{L}_Q^{2,2}$, which contains two monadic predicates F and G and two individual constants a and b.
- In $\mathcal{L}_Q^{2,2}$, we can describe 16 states, using the 16 *state descriptions* of $\mathcal{L}_Q^{2,2}$:

$Fa \,\&\, Ga \,\&\, Fb \,\&\, Gb$	$Fa \,\&\, Ga \,\&\, Fb \,\&\, {\sim}Gb$	$Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, Gb$
$Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb$	$Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, Gb$	$Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, {\sim}Gb$
$Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, Gb$	$Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb$	${\sim}Fa \,\&\, Ga \,\&\, Fb \,\&\, Gb$
${\sim}Fa \,\&\, Ga \,\&\, Fb \,\&\, {\sim}Gb$	${\sim}Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, Gb$	${\sim}Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb$
${\sim}Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, Gb$	${\sim}Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, {\sim}Gb$	${\sim}Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, Gb$
	${\sim}Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb$	

Carnapian Monadic Predicate Logical Probability 2

- Following (PI), Carnap's first measure function $\mathfrak{m}^\dagger$ assigns *equal probability* to each state description s_i of $\mathcal{L}_Q^{m,n}$. In our $\mathcal{L}_Q^{2,2}$, $\mathfrak{m}^\dagger(s_i) = \frac{1}{16}$.
- We extend $\mathfrak{m}^\dagger$ to all $p \in \mathcal{L}_Q^{m,n}$ in the standard way (the Pr of a disjunction of mutually exclusive sentences is the sum of the Pr's of its disjuncts).
- Since every $p \in \mathcal{L}_Q^{m,n}$ is equivalent to some disjunction of state descriptions, and every pair of state descriptions is mutually exclusive, this gives us a complete unconditional Pr-function $\Pr^\dagger(\cdot)$ over $\mathcal{L}_Q^{m,n}$.
- Finally, we define the conditional probability function $\Pr^\dagger(q \mid p)$ over pairs of sentences in $\mathcal{L}_Q^{m,n}$ (in the standard way) as the following *ratio*: $\frac{\Pr^\dagger(p \& q)}{\Pr^\dagger(p)}$.
- Claims of the form ⌜$\Pr^\dagger(q \mid p) = x$⌝ are *analytic in* $\mathcal{L}_Q^{m,n}$ since their truth-values are determined solely by the syntactical structure of $\mathcal{L}_Q^{m,n}$.
- But, why is *this* choice of measure function $\mathfrak{m}^\dagger$ *logical*? Logicality is ensured by the application of (PI) to state descriptions. Or is it?

Carnapian Monadic Predicate Logical Probability 3

- As it turns out, Carnap ultimately *rejects* the measure function $\mathfrak{m}^\dagger$ in favor of an alternative measure function $\mathfrak{m}^*$, for epistemic-sounding reasons.
- Carnap notes that $\mathfrak{m}^\dagger$ causes $\Pr^\dagger$ to have the following property ($b \neq a$):

$$(*) \qquad \Pr^\dagger(Fb \mid Fa) = \frac{\Pr^\dagger(Fb \,\&\, Fa)}{\Pr^\dagger(Fa)} = \frac{4 \cdot \frac{1}{16}}{8 \cdot \frac{1}{16}} = \frac{1}{2} = 8 \cdot \frac{1}{16} = \Pr^\dagger(Fb)$$

- In other words, (*) says that one object a's having property F can never raise the probability that another object b also has F. And, this *generalizes* to *any* number of Fs: $\Pr^\dagger(Fa_1 \mid Fa_2 \,\&\, \cdots \,\&\, Fa_m) = \Pr^\dagger(Fa_1)$.
- Carnap (*et al*) characterize this (in *epistemic* terms) as $\mathfrak{m}^\dagger$ leading to a function $\Pr^\dagger$ that fails to allow for "learning from experience". This is *epistemic and diachronic*, since it assumes *learning-by-conditionalization*.
- Carnap views this consequence of applying (PI) to the state descriptions of $\mathcal{L}_Q^{m,n}$ as unacceptable. Does he then think that $K_\top$ "favors" some state descriptions over others. But, which ones? Enter the $\mathfrak{m}^*$ measure …

Carnapian Monadic Predicate Logical Probability 4

- Two state descriptions s_i and s_j in $\mathcal{L}_Q^{m,n}$ are *permutations* of each other if one can be obtained from the other by a permutation of constants.
- "$Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, Gb$" can be obtained from "${\sim}Fa \,\&\, Ga \,\&\, Fb \,\&\, {\sim}Gb$" by permuting "$a$" and "$b$". Thus, "$Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, Gb$" and "${\sim}Fa \,\&\, Ga \,\&\, Fb \,\&\, {\sim}Gb$" are permutations of each other (in $\mathcal{L}_Q^{2,2}$).
- A *structure description* is a disjunction of state descriptions, each of which is a permutation of the others. $\mathcal{L}_Q^{2,2}$ has 10 structure descriptions:

$(Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, Gb) \vee ({\sim}Fa \,\&\, Ga \,\&\, Fb \,\&\, {\sim}Gb)$
$(Fa \,\&\, Ga \,\&\, Fb \,\&\, {\sim}Gb) \vee (Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, Gb)$
$(Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, Gb) \vee ({\sim}Fa \,\&\, Ga \,\&\, Fb \,\&\, Gb)$
$(Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb) \vee ({\sim}Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, Gb)$
$(Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb) \vee ({\sim}Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, {\sim}Gb)$
$({\sim}Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb) \vee ({\sim}Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, Gb)$
$Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, {\sim}Gb$
$Fa \,\&\, Ga \,\&\, Fb \,\&\, Gb$
${\sim}Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, Gb$
${\sim}Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb$

- The measure $\mathfrak{m}^*$ [$\therefore \Pr^*$] assigns equal probability to *structure descriptions*. $\mathfrak{m}^*$ takes individuals to be *indistinguishable* (analogy: Bose-Einsten statistics in QM), and $\mathfrak{m}^\dagger$ does not (analogy: Fermi-Dirac statistics in QM). Otherwise, they operate in the same way.

Carnapian Monadic Predicate Logical Probability 5

- We can then define $\Pr^*(\bullet)$ and $\Pr^*(\bullet \mid \bullet)$ in terms of $\mathfrak{m}^*$, by assuming equiprobability of *states within* structure descriptions. Let's compare the $\mathfrak{m}^\dagger$ and $\mathfrak{m}^*$ distributions over the 16 state descriptions of $\mathcal{L}_Q^{2,2}$:

Fa	Ga	Fb	Gb	State Descriptions (s_i)	$\mathfrak{m}^\dagger(s_i)$	$\mathfrak{m}^*(s_i)$
T	T	T	T	$Fa \,\&\, Ga \,\&\, Fb \,\&\, Gb$	1/16	1/10
T	T	T	F	$Fa \,\&\, Ga \,\&\, Fb \,\&\, {\sim}Gb$	1/16	1/20
T	T	F	T	$Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, Gb$	1/16	1/20
T	T	F	F	$Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb$	1/16	1/20
T	F	T	T	$Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, Gb$	1/16	1/20
T	F	T	F	$Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, {\sim}Gb$	1/16	1/10
T	F	F	T	$Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, Gb$	1/16	1/20
T	F	F	F	$Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb$	1/16	1/20
F	T	T	T	${\sim}Fa \,\&\, Ga \,\&\, Fb \,\&\, Gb$	1/16	1/20
F	T	T	F	${\sim}Fa \,\&\, Ga \,\&\, Fb \,\&\, {\sim}Gb$	1/16	1/20
F	T	F	T	${\sim}Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, Gb$	1/16	1/10
F	T	F	F	${\sim}Fa \,\&\, Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb$	1/16	1/20
F	F	T	T	${\sim}Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, Gb$	1/16	1/20
F	F	T	F	${\sim}Fa \,\&\, {\sim}Ga \,\&\, Fb \,\&\, {\sim}Gb$	1/16	1/20
F	F	F	T	${\sim}Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, Gb$	1/16	1/20
F	F	F	F	${\sim}Fa \,\&\, {\sim}Ga \,\&\, {\sim}Fb \,\&\, {\sim}Gb$	1/16	1/10

- Now, $\mathfrak{m}^*$ does *not* have the "no learning from experience" property (*):

$$\Pr^*(Fb \mid Fa) = \frac{\Pr^*(Fb \,\&\, Fa)}{\Pr^*(Fa)} = \frac{2 \cdot \frac{1}{10} + 2 \cdot \frac{1}{20}}{2 \cdot \frac{1}{10} + 6 \cdot \frac{1}{20}} = \frac{3}{5} > 2 \cdot \frac{1}{10} + 6 \cdot \frac{1}{20} = \frac{1}{2} = \Pr^*(Fb)$$

- Generally, $\Pr^*(\bullet \mid \bullet)$ says that the more objects that are assumed to have F, the more probable it is that other objects will also have F. This is called *instantial relevance*. Carnap prefers $\Pr^*$ over $\Pr^\dagger$ for this reason.
- Is Carnap committed to the view that $K_\top$ favors certain state descriptions over others? He prefers $\mathfrak{m}^*(Fa \,\&\, Ga \,\&\, Fb \,\&\, Gb) > \mathfrak{m}^*(Fa \,\&\, Ga \,\&\, Fb \,\&\, {\sim}Gb)$ to $\mathfrak{m}^\dagger(Fa \,\&\, Ga \,\&\, Fb \,\&\, Gb) = \mathfrak{m}^\dagger(Fa \,\&\, Ga \,\&\, Fb \,\&\, {\sim}Gb)$. But, *why*?
- Carnap realizes that if a theory of "logical probability" is to provide a foundation for claims about *evidence*, then it *must* be able to furnish Pr-models with *correlations* between *logically independent* claims.
- Moral: "logical" probability models such as these are unable to emulate a broad enough range of probability functions, so as ($\mathcal{D}_3$) to be generally *applicable* for modeling evidential relations in epistemic contexts.
- Indeed, there seems to be some wrongheaded about trying to "formalize inductive inference" (or relations of evidential support) in the first place.

A Closer Look at Inductive Logic and Applicability 1

- Carnap suggested various *bridge principles* for connecting inductive logic and inductive epistemology. The most well-known of these was:
  - **The Requirement of Total Evidence**. In the application of inductive logic to a given knowledge situation, the total evidence available must be taken as a basis for determining the degree of evidential support.
- A more precise way of putting this principle is:

(RTE) E evidentially supports H *for an agent* S *in an epistemic context* C $\iff \mathfrak{c}(H, E \mid K) > r$, where K is S's *total evidence in* C.

- For Carnap, $\mathfrak{c}(H, E \mid K) = \Pr_{\mathsf{T}}(H \mid E \,\&\, K)$, where $\Pr_{\mathsf{T}}$ is a suitable "logical" probability function. So, we can restate *Carnap's* (RTE) as follows:

(RTE_C) E evidentially supports H *for an agent* S *in an epistemic context* C $\iff \Pr_{\mathsf{T}}(H \mid E \,\&\, K) > r$, where K is S's *total evidence in* C.

- Carnap's version of (RTE) faces a challenge (first articulated by Popper) involving *probabilistic relevance vs high conditional probability*.

A Closer Look at Inductive Logic and Applicability 2

- Popper discusses examples of the following kind (we've seen an example like this in the class before), which involve testing for a rare disease.
  - Let E report a positive test result for a very rare disease (for someone named John), and let H be the (null) hypothesis that John does *not* have the disease in question. We assume further that John knows (his K entails) the test is highly reliable, and that the disease is very rare.
- In such an example, it is plausible (and Carnap should agree) that (to the extent that $\Pr_{\mathsf{T}}$ is *applicable* to modeling the *epistemic* relations here):

(1) $\Pr_{\mathsf{T}}(H \mid E \,\&\, K)$ is very high.

(2) But, $\Pr_{\mathsf{T}}(H \mid E \,\&\, K) < \Pr_{\mathsf{T}}(H \mid K)$.

- Because of (2), it would be odd to say that E *supports* H (for John) in this context. (2) suggests that E is (intuitively) evidence *against* H here.
- But, because of (1), Carnap's (RTE_C) implies that E supports H (for John) here. This looks like a counterexample to [the $\Leftarrow$ of] Carnap's (RTE_C).

A Closer Look at Inductive Logic and Applicability 3

- This suggests the following refinement of Carnap's (RTE_C):

(RTE'_C) E evidentially supports H *for an agent* S *in an epistemic context* C $\implies \Pr_{\mathsf{T}}(H \mid E \,\&\, K) > \Pr_{\mathsf{T}}(H \mid K)$, where K is S's *total evidence in* C.

- In other words, (RTE'_C) says that *evidential support in (for* S *in* C*) implies probabilistic relevance, conditional upon* K (for a suitable $\Pr_{\mathsf{T}}$ function).
- Note: this only states a *necessary* condition for evidential support.
- While (RTE'_C) avoids Popper's objection, it faces serious challenges of its own (e.g., Goodman's "Grue" example — more on that later). Here's one:
- Consider any context in which S already knows (*with certainty*) that E is true. That is, S's total evidence in the context *entails* E ($K \vDash E$).
- In such a case, $\Pr(H \mid E \,\&\, K) = \Pr(H \mid K)$, for ***any*** probability function Pr. Thus, (RTE'_C) implies that E *cannot support anything* (for S, in any such C). This shows that (RTE'_C) isn't a correct principle either. ["Old Evidence"]

A Closer Look at Inductive Logic and Applicability 4

- I think this whole way of approaching inductive logic is wrongheaded.
- First, why must Pr *itself* be logical, if $\mathfrak{c}$ (which is defined in terms of Pr) is to be logical? Analogy: must the truth-value assignment function v *itself* be logical, if $\vDash$ (which is defined in terms of v) is to be logical?
  - But: there is a crucial disanalogy here, which I will discuss below.
- Second, Carnap's proposal $\mathfrak{c}(H, E \mid K) = \Pr_{\mathsf{T}}(H \mid E \,\&\, K)$ is suspect, because (as Popper pointed out) it is not sensitive to *probabilistic relevance*.
  - Note: this undermines Carnap's argument for the "logicality" of (PI).
- Third, the applicability desideratum ($\mathcal{D}_3$) may be fundamentally misguided. The search for logic/epistemology "bridge principles" is fraught with danger, even in the *deductive* case. And, since IL is supposed to *generalize* DL, it will also face these dangers *and new ones* (as above).
  - I think this is the true (but, surprisingly, un-appreciated) lesson of Goodman's "grue" example. [See my "grue" paper for in-depth discussion.]

An Alternative Conception of Inductive Logic 1

- In light of the above considerations, we might seek a measure $\mathfrak{c}$ satisfying the following (provided that E, H, and K are *logically contingent*):

$$\mathfrak{c}(H, E \mid K) \text{ should be } \begin{cases} \text{Maximal } (> 0, \text{ constant}) & \Leftarrow E \,\&\, K \vDash H. \\ > 0 \text{ (confirmation rel. to } K) & \Rightarrow \Pr(H \mid E \,\&\, K) > \Pr(H \mid K). \\ = 0 \text{ (irrelevance rel. to } K) & \Rightarrow \Pr(H \mid E \,\&\, K) = \Pr(H \mid K). \\ < 0 \text{ (disconfirmation rel. to } K) & \Rightarrow \Pr(H \mid E \,\&\, K) < \Pr(H \mid K). \\ \text{Minimal } (< 0, \text{ constant}) & \Leftarrow E \,\&\, K \vDash \sim H. \end{cases}$$

- Carnap would add: "and Pr should be a 'logical' probability function $\Pr_{\mathsf{T}}$". But, I suggested that this was a mistake. OK, but then what do *I* say about the Pr's above? There is an implicit quantifier over the Pr's above...
  - $\exists$ is *too weak* a quantifier here, since there will *always* be *some* such Pr.
  - $\forall$ is *too strong* a quantifier here, because that is *demonstrably false*!
  - What's the alternative? The alternative is that Pr is a *parameter* in $\mathfrak{c}$ itself. That is, perhaps Pr is simply an *argument* of the function $\mathfrak{c}$.

An Alternative Conception of Inductive Logic 2

- Here's the idea. Confirmation is a *four*-place relation, between E, H, K, *and a probability function* Pr. The resulting relation is still *logical* in Carnap's sense, since, *given* a choice of Pr, $\mathfrak{c}$ is logically (mathematically, if you prefer) determined, provided only that $\mathfrak{c}$ is defined in terms of Pr.
- So, on this conception, desiderata ($\mathcal{D}_1$) and ($\mathcal{D}_2$) are satisfied.
- As usual, the subtle questions involve the applicability desideratum ($\mathcal{D}_3$).
- What do we say about that? Well, I think any naive "bridge principle" like (RTE or RTE′) is doomed to failure. But, perhaps there is *some* connection.
- Thinking back to the deductive case, there may be *some* connection between deductive *logic* and deductive *inference*. But, what is it?
- This is notoriously difficult to say. The best hope seems to be that there is *some* connection between *knowledge* and entailment. [Note: connecting or bridging *justified belief* and *entailment* seems much more difficult.]

Inductive–Statistical Explanation II

- The simplest schema for an I–S explanation would be:

$$\begin{array}{l} \Pr(Gx \mid Fx) = r \\ Fb \\ \hline\hline Gb \end{array} \qquad [r]$$

- Here, "$\Pr(Gx \mid Fx) = r$" is a 'statistical law' which says that the relative frequency of Gs among Fs is r. Note: this is a *non-modal* frequency, just as $(\forall x)(Fx \supset Gx)$ is a non-modal (extensional) "universal law".
- "$[r]$" is supposed to be the argument's *inductive strength*. But, is it?
- **Example**: John Jones (b) recovers quickly (Gb) from Strep (Fb). Most strep infections (Fx) clear up quickly (Gx) when treated with penicillin (Hx). Thus, we have the following I–S explanation of the fact that Gb:

$$\begin{array}{l} \Pr(Gx \mid Fx \,\&\, Hx) = r \approx 1 \\ Fb \,\&\, Gb \\ \hline\hline Gb \end{array} \qquad [r \approx 1]$$

Inductive–Statistical Explanation III

- But, what if we were to subsequently learn that John Jones was infected with a *penicillin-resistant* strain of Strep (Jb)? Plausibly, this would lead to a "strong" I–S explanation of $\sim Gb$, as follows:

$$\begin{array}{l} \Pr(\sim Gx \mid Fx \,\&\, Hx \,\&\, Jx) = r_1 \approx 1 \\ Fb \,\&\, Gb \,\&\, Jb \\ \hline\hline \sim Gb \end{array} \qquad [r_1 \approx 1]$$

- This is the *reference class problem* for frequency interpretations of probability. Depending on which reference class we decide to include b in, we get different "$\Pr(Gb)$"s. We can have "strong arguments" for both Gb and $\sim Gb$! This is *impossible* with deduction (and D–N).
- In confirmation theory, one traditionally adds the requirement of *total evidence — viz.*, the requirement that *no additional evidence that would change "the* $\Pr(Gb)$*" is available in the context*. This will *not* work here, since the *explanandum* (say, Gb) is *known*! What shall we do?

Inductive–Statistical Explanation IV

- Hempel adds the *requirement of maximal specificity* (RMS). Here, we assume that **P** is the conjunction of all of the premises of the I–S explanation, and K is the available background knowledge.

(RMS) If **P** & K implies that b belongs to a class F_1 and that F_1 is a subclass of F, then **P** & K must also imply a statement specifying the statistical probability (frequency) of G in F_1, say

$$\Pr(G \mid F_1) = r_1.$$

And, $r_1 = r$, unless the probability statement in question is simply a theorem of probability calculus (*e.g.*, $\Pr(G \mid F_1 \,\&\, G) = 1$).

- The "unless" clause in (RMS) is there to block the trivialization which would otherwise arise from the fact that we know from K that b is both F and G (since probability calculus implies that $\Pr(Gx \mid Fx \,\&\, Gx) = 1$).
- Note, deductively valid arguments *vacuously* satisfy (RMS). *Why*?

Inductive–Statistical Explanation V

- According to Hempel, the *non-monotonicity* of inductive inferences leads, inevitably, to the *epistemic relativity of statistical explanation*:

  > "The concept of statistical explanation for particular events is essentially relative to a given knowledge situation as represented by a class K of accepted statements."

- It seems clear that the concept of inductive strength is (in some sense) "relativistic", but why say that it is *epistemically* "relativistic"?
  - We don't take this attitude toward special relativity (*vs* Newtonian theory). That is, we *do* say that, *e.g.*, "simultaneity" is (in some sense) *relativistic* (*i.e.*, that it is *relative to **a frame of reference***), but we *don't* say that it is *relative to **what we know***. Why the different attitude when it comes to inductive logic (*vs* deductive logic)? This is a deep question about deductive *vs* inductive logic (which is, alas, beyond the scope of this seminar).
- Alberto Coffa's discussion (as quoted Salmon) is quite illuminating on this issue of "epistemic relativity". He says, about "confirmation", that:

> Although the syntactic form of expressions like "hypothesis h is well-confirmed" may mislead us into believing that confirmation is a property of sentences, closer inspection reveals the fact that it is a relation between sentences and knowledge situations and that the concept of confirmation cannot be properly defined . . . without reference to sentences intended to describe a knowledge situation.

But, interestingly, Coffa sings a different tune about "explanation":

> . . . the possibility of a notion of true explanation . . . is not just a desirable but ultimately dispensable feature of a model of explanation: it is the *sine qua non* of its realistic, non-psychologistic inspiration. It is because certain features of the world can be deterministically responsible for others that we can describe a concept of true deductive explanation . . . If there are features of the world which can be non-deterministically responsible for others, then we should be able to define a model of true inductive explanation.

- Why not say the same thing about inductive *arguments/logic*? Isn't the notion of a "valid argument" the *sine qua non* of the realistic, non-psychologistic inspiration of *logic* (inductive or otherwise)?
- OK, enough of that. I'll be doing a seminar on *confirmation next year* . . .
  - Preview: traditionally, it is just assumed that inductive logic *is* (*intrinsically*) *epistemic*. But, this is *not* (traditionally) assumed about deductive logic.

Inductive–Statistical Explanation VI

- You may already be able to guess what some of the problems with I–S explanation are going to be. Traditional accounts of " inductive strength" are plagued by their own problems of *relevance*.
- "$\Pr(X \mid Y)$ is high" is neither necessary nor sufficient for "Y is relevant to X". The Fred Fox example, and the paresis/syphilis example (above) show that *relevance* is an important aspect of explanations.
- Both the D–N and the I–S accounts suffer from ***problems of relevance***.
  [Kaplanian & hexed salt examples show D–N is has *several* such problems.]
- While high probability is neither necessary nor sufficient for explanatory power (because it is detached from *relevance*), it may still be plausible that (*other things being equal — e.g.*, assuming that there *is* relevance, *etc.*) high*er* probabilities tend to lead to *better* explanations.
  - Michael Strevens discusses this in his paper "Do Large Probabilities Explain Better?", which is posted on the website. He will be our guest on 4/12.

Statistical–Relevance Explanation I

- Salmon, Greeno, Jeffrey, and others were among the first to question the high probability requirement of the I–S account. Jeffrey says:

  Consider a genuinely indeterministic coin which is biased strongly (p = 0.9) toward heads when tossed. Suppose that if it is not tossed the coin has probability of 0.5 of being in either the heads or tails position and that whether or not the coin is tossed is the only factor that is statistically relevant to whether it is heads or tails. According to the I–S model, if the coin is tossed and comes up heads, we can explain this outcome by appealing to the fact that the coin was tossed (since under this condition the probability of heads is high) but if the coin is tossed and comes up tails we cannot explain this outcome, since its probability is low ... The fact that the coin has been tossed is the only factor relevant to either outcome and that factor is common to both outcomes — *once we cite the toss ... we leave nothing out that influences the outcome.*

- There are various worries one might have about this line of argument...

- First, as Woodward points out:

  such arguments presuppose that it is not possible for all of the information that is relevant to some (event) E to be insufficient to explain it. It is far from self-evident that this presupposition is correct.

- Second, as we will see shortly, the theories that came after I–S tended to appeal to *statistical* (or *probabilistic*) *relevance* relations. And, it's not clear whether that is essential to Jeffrey's line of argument. [Even if there isn't statistical relevance, Jeffrey's presupposition might (still) kick-in, since we might (still) have "left nothing out that influences the outcome."]
- Salmon and Greeno formulated theories in which the key probabilistic fact is a fact about probabilistic *relevance* — *not high* (posterior) *probability* of the explanandum (given the explanans, *etc.*).
- Accounts involving *statistical relevance* as the key attribute face special problems of their own — problems not faced by the I–S account.
- Perhaps the most important of these involves what is known as *Simpson's Paradox*. Cartwright discusses a nice example...

Statistical–Relevance Explanation II

- In the early 80's there was a positive correlation between being female (F) and being rejected from Berkeley's graduate school (R).
- This (initially) raised some suspicions about the possibility of sexual discrimination in the admissions process for Berkeley's grad school.
- Symbolically, $\Pr(R \mid F) > \Pr(R)$. That is, being female was *statistically relevant* to being rejected from Berkeley grad school.
- *But*, if we *partition* the applicants according to the department to which they applied: $\{D_1, \ldots, D_n\}$, then *all such correlations disappear*!
- That is, $\Pr(R \mid F \,\&\, D_i) = \Pr(R \mid D_i)$, for *all* i.
- Should we still be suspicious about sexual discrimination? Or, more relevantly here, should we still think that gender (of an applicant to BGS in the 80's) was *explanatorily relevant* to rejection (or acceptance)?
  - Note: at this point (especially), the *type/token distinction* becomes especially salient. We'll discuss that issue (a lot) when we get to Woodward.

Statistical–Relevance Explanation III

- Simpson's Paradox forces the statistical relevance theorist to make some special maneuvers. Enter the notion of "homogeneous partition".
- Salmon's original S–R account is intended to provide an answer to the question "Why does this (member of the reference class) A have the attribute B?" [If the question is not stated this precisely, then Salmon suggests using "pragmatics" to determine the reference class.]
- An S–R explanation (of why this A is a B), consists of the (prior) probability of B (given A), a homogeneous relevant partition $\{A \,\&\, C_i\}$ of A with respect to B, the (posterior) probabilities of B in each member $A \,\&\, C_i$ of the partition, and a statement of the location of the individual in question in a particular cell $A \,\&\, C_k$ of the partition. That is:
  - $\Pr(B \mid A) = p$
  - $\Pr(B \mid A \,\&\, C_i) = p_i$
  - $\{A \,\&\, C_i\}$ is a homogenous relevant partition of A with respect to B
  - b is a member of $A \,\&\, C_k$

Statistical–Relevance Explanation IV

- To fix ideas, let's return to the Berkeley graduate school example. Let b be some applicant (A) who was rejected (B). And, we want to know why this applicant (b) was rejected (*i.e.*, why this A (b) is a B).
- A *partition* $\{A \,\&\, C_i\}$ of a class A is a collection of mutually exclusive and exhaustive subsets of A. Each subclass $A \,\&\, C_i$ in the partition $\{A \,\&\, C_i\}$ is called a *cell* of the partition.
- A partition $\{A \,\&\, C_i\}$ of A is *relevant* with respect to B if the probability (*i.e.*, the relative statistical frequency!) of B in each cell $A \,\&\, C_i$ of the partition is different from each other cell — that is, if:

$$\Pr(B \mid A \,\&\, C_i) \neq \Pr(B \mid A \,\&\, C_j), \text{ all } i \neq j.$$

- A partition F is *homogeneous* with respect to B if *no relevant partition (wrt B) can be made within F*. *Objectively* homogeneous = no relevant partition can be made *in principle*; *epistemically* homogeneous = no relevant partition is *known* (presumably, by the explain*er*).

- In this example, we have a relevant partition of A with respect to B:
- That partition is $\{A \,\&\, C_i\}$: the partition in which the C_i are the *genders* of the applicants. If C_1 = Male, and C_2 = Female, then:

$$\Pr(B \mid A) = p,\ \Pr(B \mid A \,\&\, C_1) = p_1,\ \Pr(B \mid A \,\&\, C_2) = p_2$$

  Here, $p_2 < p < p_1$. Therefore, the partition $\{A \,\&\, C_i\}$ is *relevant* to B.
- Intuitively, the partition $\{A \,\&\, C_i\}$ is *not* "explanatory" wrt B, because there is a *further* set of (intuitively) *relevant factors* $\{D_i\}$ such that:

$$\Pr(B \mid A \,\&\, C_1 \,\&\, D_i) = \Pr(B \mid A \,\&\, C_2 \,\&\, D_i), \text{ for all } i$$

- Does this mean that the original partition $\{A \,\&\, C_i\}$ is *not homogeneous* with respect to B? If so, that would undermine an explanation of Ab using the $\{A \,\&\, C_i\}$ partition (*i.e.*, an explanation in terms of *gender*).
- We can't tell from the information so far whether the finer-grained partition $\{A \,\&\, C_i \,\&\, D_j\}$ is relevant. For that, we'd need to check and see whether, *e.g.*, $\Pr(B \mid A \,\&\, C_2 \,\&\, D_i) \neq \Pr(B \mid A \,\&\, C_2 \,\&\, D_j)$, for $i \neq j$.
- Here, the answer happens to be YES. A vindication of the S–R model?

Statistical–Relevance Explanation V

- Intuitively, "homogeneity" is *intended* to *block* (*e.g.*) the explanatoriness of gender for acceptance in the Berkeley grad school example. And, it does seem to do the trick — *in this case*. But, will it *always* work?
- If $\Pr(\cdot \mid \cdot)$ is just *statistical frequency*, then there may *happen to be* further "relevant" partitions of the (actual) data — in Salmon's sense. Or, there may *happen to be* no further "relevant" partitions — in Salmon's sense. Does "S–R-relevance" track *explanatory* relevance?
- Fact about The Berkeley Case: partitioning *further within* $\{A \,\&\, C_i \,\&\, D_j\}$, according to $\{Z_1, Z_2\} \stackrel{\text{def}}{=}$ the first letter of the applicant's surname is (is not) between "F" and "K", yields a "*statistically relevant* partition".
- Do we want to say that $\{A \,\&\, C_i \,\&\, D_j \,\&\, Z_k\}$ is an *explanatorily* relevant partition, and/or that the existence of this finer-grained partition *undermines* the explanatory probative value of $\{A \,\&\, C_i \,\&\, D_j\}$?
- What we seem to need is "homogeneity" wrt "all *explanatorily* relevant factors". Idea: *causal* relevance? This is a nice lead-in to Woodward …