**Inverse Probability and Bayes's Theorem II**

- Here's a famous example, illustrating the subtlety of Bayes's Theorem:
  - The (unconditional) probability of breast cancer is 1% for a woman at age forty who participates in routine screening. The probability of such a woman having a positive mammogram, given that she has breast cancer, is 80%. The probability of such a woman having a positive mammogram, given that she does not have breast cancer, is 10%. What is the probability that such a woman has breast cancer, given that she has had a positive mammogram in routine screening?

- We can formalize this, as follows. Let $H$ = such a woman (age 40 who participates in routine screening) has breast cancer, and $E$ = such a woman has had a positive mammogram in routine screening. Then:
  $$\Pr(E \mid H) = 0.8, \Pr(E \mid \neg H) = 0.1, \text{ and } \Pr(H) = 0.01.$$  

- Question (like Hacking's O.Q. #5): What is $\Pr(H \mid E)$? What would you guess? Most experts guess a pretty high number (near 0.8, usually).

**Inverse Probability and Bayes's Theorem III**

- Hacking's O.Q. #6: You are a physician. You think it is quite probable (say 90% probable) that one of your patients has strep throat ($S$). You take some swabs from the throat and send them to the lab for testing. The test is imperfect, with the following likelihoods ($Y$ is + result, $N$ is −):
  - $\Pr(Y \mid S) = 0.7, \Pr(Y \mid \neg S) = 0.1$
  - You send five successive swabs to the lab, from the same patient. You get the following results, in order: $Y, N, Y, N, Y$. What is $\Pr(S \mid YNYNY)$?
  - Hacking: Assume that the 5 test results are conditionally independent, given both $S$ and $\neg S$, i.e., that $S$ screens-off the 5 tests results. So:
    - $\Pr(YNYNY \mid S) = 0.7 \cdot 0.3 \cdot 0.7 \cdot 0.3 \cdot 0.7 = 0.03087$
    - $\Pr(YNYNY \mid \neg S) = 0.1 \cdot 0.9 \cdot 0.1 \cdot 0.9 \cdot 0.1 = 0.00081$
    - $\Pr(S \mid YNYNY) = \frac{\Pr(YNYNY \mid S) \cdot \Pr(S)}{\Pr(YNYNY \mid S) \cdot \Pr(S) + \Pr(YNYNY \mid \neg S) \cdot \Pr(\neg S)}$
      - $\Pr(S \mid YNYNY) = \frac{0.03087 \cdot 0.7}{0.03087 \cdot 0.7 + 0.00081 \cdot 0.1} = 0.997$
General Analysis of Hacking’s “Odd Question #6”

- If $n$ is the number of Y results, then $(5 - n)$ is the number of N results (out of 5 results). Bayes’s theorem allows us to calculate $\text{Pr}(S | E_n)$, where $E_n$ is evidence consisting of $n$ Y results and $(5 - n)$ N results (any order):

$$\frac{\text{Pr}(E_n | S) \cdot \text{Pr}(S)}{\text{Pr}(E_n | S) \cdot \text{Pr}(S) + \text{Pr}(E_n | \sim S) \cdot \text{Pr}(\sim S)} = \frac{0.7^n \cdot 0.3^{5-n} \cdot 0.9}{0.7^n \cdot 0.3^{5-n} \cdot 0.9 + 0.1^n \cdot 0.95^{5-n} \cdot 0.1}$$

---

**An Anecdotal Prelude to “Interpretations” of Probability**

- After the O.J. trial, Alan Dershowitz remarked that “fewer than 1 in 1,000 women who are abused by their mates go on to be killed by them”.

- He said “the probability” that Nicole Brown Simpson (N.B.S.) was killed by her mate (O.J.) — *given that he abused her* — was less than 1 in 1,000.

- Presumably, this was supposed to have some consequences for people’s *degrees of confidence* (degrees of belief) in the hypothesis of O.J.’s guilt.

- The debate that ensued provides a nice segue from our discussion of the formal theory of probability calculus to its “interpretation(s)”.

- Let $A$ be the proposition that N.B.S. is abused by her mate (O.J.), let $K$ be the proposition that N.B.S. is killed by her mate (O.J.), and let $\text{Pr}($·$)$ be whatever probability function Dershowitz has in mind here, over the salient algebra of propositions. Dershowitz is saying the following:

$$\text{(1)} \quad \text{Pr}(K | A) < \frac{1}{1000}$$

---

- Shortly after Dershowitz’s remark, the statistician I.J. Good wrote a brief response in *Nature*. Good pointed out that, while Dershowitz’s claim may be true, it is not salient to the case at hand, since it ignores evidence.

- Good argues that what’s relevant here is is the probability that she was killed by O.J., given that she was abused by O.J. and *that she was killed*.

- After all, we do know that Nicole was killed, and (plausibly) this information should be taken into account in our probabilistic musings.

- To wit: let $K'$ be the proposition that N.B.S was killed (by someone). Using Dershowitz’s (1) as a starting point, Good does some *ex cathedra* back-of-the-envelope calculations, and he comes up with the following:

$$\text{(2)} \quad \text{Pr}(K | A & K') \approx \frac{\frac{1}{2}}{\frac{1}{5}} \gg \frac{1}{1000}$$

- This would seem to make it far more probable that O.J. is the killer than Dershowitz’s claim would have us believe. Using statistical data about murders committed in 1992, Merz & Caulkins “estimated” that:

$$\text{(3)} \quad \text{Pr}(K | A & K') \approx \frac{4}{7}$$

---

- This would seem to provide us with an *even greater* “estimate” of “the probability” that N.B.S. was killed by O.J. Dershowitz replied to analyses like those of Good and Merz & Caulkins with the following rejoinder:

… whenever a woman is murdered, it is highly likely that her husband or her boyfriend is the murderer without regard to whether battery preceded the murder. The key question is how salient a characteristic is the battery as compared with the relationship itself. Without that information, the 80 percent figure [as in Merz & Caulkins’ estimation] is meaningless. I would expect that a couple of statisticians would have spotted this fallacy.

- Dershowitz’s rejoinder seems to trade on something like the following:

$$\text{(4)} \quad \text{Pr}(K | K') \approx \text{Pr}(K | A & K') \quad \text{[i.e.,} \ K', \text{not} \ A, \text{is doing the real work here]}$$

- Not to be outdone, Merz & Caulkins give the following “estimate” of the salient probabilities (again, this is based on statistics for 1992):

$$\text{(5)} \quad \text{Pr}(K | K') \approx 0.29 \ll \text{Pr}(K | A & K') \approx 0.8$$

- We could continue this dialectic *ad nauseam*. I’ll stop here. This anecdote raises several key issues about “interpretations” and “applications” of Pr.
- Our discussants seem to be talking about some kind of "objective" probabilities involving N.B.S.'s murder (and murderer) in particular.

- But, the "estimates" Merz & Caulkins appeal to involve statistical frequencies of murders in some population.

- First, are there such things as objective probabilities at all? If so, what are they (are there different kinds?) and what determines them?

- More specifically, are there objective probabilities of token events, or only frequencies in populations? If there are such probabilities of token events, then how (if at all) do they relate to frequencies?

- Specifically, which population is "the right one" in which to include the token event (this is known as the reference class problem)?

- Finally, how are "objective" probabilities related to degrees of belief.

- Generally, how are "objective" and "subjective" probabilities related?

- We'll be thinking more about some of these questions in the next unit.

- But, first, we're going to back-up and start "from scratch" …

---

**T-on-I : Truth :: Pr-on-M : Probability (I)**

- In logic (and formal semantics), we have a formal notion called truth-on-an-interpretation (or truth-on-I). This is not truth (simpliciter).

- It's useful to think about examples now. Here's a very simple example.

- Consider a 2-atom sentential language \( L \), where the atoms are extra-systematically understood as having the following content:
  - \( X \) \( \equiv \) John is unmarried.
  - \( Y \) \( \equiv \) John is a bachelor.

- As usual, we can picture all four interpretations of \( L \), as follows:

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
<th>Interpretations</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>( I_1 )</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>( I_2 )</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>( I_3 )</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>( I_4 )</td>
</tr>
</tbody>
</table>

- Facts about truth-on-\( I_1 \) do not depend on (extra-systematic) content.

---

**T-on-I : Truth :: Pr-on-M : Probability (II)**

- Specifically, we have the following facts about truth-on-\( I_1 \):
  - \( X \) is T-on-\( I_1 \) and T-on-\( I_2 \), but \( X \) is F-on-\( I_3 \) and F-on-\( I_4 \).
  - \( Y \) is T-on-\( I_1 \) and T-on-\( I_3 \), but \( X \) is F-on-\( I_2 \) and F-on-\( I_4 \).

- Indeed, all facts about truth-on-\( I_1 \) are determined for all sentences \( p \) of \( L \) — just by our conventions about truth-tables for truth-functional logic.

- In this sense, truth-on-\( I_1 \) does not depend on the extra-sytematic content of the sentences of \( L \). But, the truth (simpliciter) of sentences does.

- In this sense, truth is external to logic (and to formal semantics).

- OK, so then what is truth, and how is it related to truth-on-\( I \)?

---

**T-on-I : Truth :: Pr-on-M : Probability (III)**

- For each interpretation \( I_i \) of \( L \), there is a corresponding state-description \( s_i \) of \( L \). As a result, \( \' p \) is T-on-\( I_i \) \) is synonymous with \( \' s_i \models p \).”

- What this reveals is that truth-on-\( I \) is a systematic logical concept.

- On the other hand, truth is an extra-systematic concept.

- In our example, \( Y \) extra-systematically entails (or conceptually necessitates) \( X \) [\( Y \not\equiv X \)], since it is a conceptual truth that all bachelors are unmarried.

- This allows us to extra-systematically rule-out the truth of the third state description \( s_3 \) of \( L \). That is, \( s_3 \) cannot be true, despite the fact that \( Y \not\equiv X \).

- This is a very strong sense of “ruling-out an interpretation”. There are also two weaker senses of “ruling out” that can obtain:
  - Although \( Y \not\equiv X \), \( Y \) conceptually probabilifies \( X \).
  - \( Y \) epistemically (but not conceptually) probabilifies \( X \).

- Next: examples of each of these other two “grades of ruling-out”.

---
Consider the following example, again involving two sentences $X$ and $Y$.
- $X \not\implies$ The coin will land heads when it is tossed.
- $Y \not\implies$ The coin is heavily biased in favor of heads.

Here, we have neither $Y \implies X$ nor $Y \iff X$. But, $I_3$ still seems somehow "inappropriate". If the coin was 2-headed, then we would have $Y \implies X$.

There is some sort of extra-systematic conceptual probability relation between $Y$ and $X$. But, $\neg X$ and $Y$ are not conceptually inconsistent here.

The natural thing to do here is to try to represent this as some sort of probabilistic extra-systematic constraint. But, which constraint is it?
1. $Y \not\implies \Pr(X) \approx 1$. [Y e.s.-entails that $X$ is highly probable. Meaningful?]
2. $\Pr(Y \rightarrow X) \approx 1$. [The conditional $Y \rightarrow X$ is highly probable.]
3. $\Pr(X \mid Y) \approx 1$. [The conditional probability of $X$, given $Y$, is high.]

None of these rules-out the truth of $s_3$. But, they all place e.s.-constraints on how probable $s_3$ is. For instance, (2) forces $\Pr(s_3)$ to be low (why?).

That leads to the following extra-systematic revision of our initial STT:

<table>
<thead>
<tr>
<th>$X$</th>
<th>$Y$</th>
<th>Interpretations/S.D.'s</th>
<th>Models ($M$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>$I_1 / s_1$</td>
<td>$a_1 \in [0, 1]$</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>$I_2 / s_2$</td>
<td>$a_2 \in [0, 1]$</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>$I_3 / s_3$</td>
<td>0</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>$I_4 / s_4$</td>
<td>$a_4 = 1 - (a_1 + a_2)$</td>
</tr>
</tbody>
</table>

In other cases, we will not be able to rule-out any interpretations. But, we will be able to rule-out certain probability assignments/models. E.g.:
- $X \not\implies$ The coin will land heads when it is tossed.
- $Y \not\implies$ The coin is heavily biased in favor of heads.

In this case, let’s assume the right constraint is $\Pr(X \mid Y) \approx 1$. Then, this will impose the following extra-systematic constraint on our initial STT:

$$\frac{a_1}{a_1 + a_3} \approx 1$$

This doesn’t rule-out any interpretations, but it does rule-out some probability models. Finally, there is a third "grade of ruling-out"...

Initially, we have only systematic constraints. Specifically, we have no systematic logical relations between atomic sentences, and the only systematic probabilistic constraints are $a_i \in [0, 1]$ and $\sum_i a_i = 1$. E.g.:

<table>
<thead>
<tr>
<th>$X$</th>
<th>$Y$</th>
<th>Interpretations/S.D.'s</th>
<th>Models ($M$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>$I_1 / s_1$</td>
<td>$a_1 \in [0, 1]$</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>$I_2 / s_2$</td>
<td>$a_2 \in [0, 1]$</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>$I_3 / s_3$</td>
<td>0</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>$I_4 / s_4$</td>
<td>$a_4 = 1 - (a_1 + a_2 + a_3)$</td>
</tr>
</tbody>
</table>

Then, we associate extra-systematic contents with the atoms, e.g.:
- $X \not\implies$ John is unmarried.
- $Y \not\implies$ John is a bachelor.

In this case, we can conceptually rule-out interpretation $I_3$ on extra-systematic grounds. In other words, $s_3$ is (necessarily) false.

Here is another example of a pair of sentences:
- $X \not\implies$ The ball is black.
- $Y \not\implies$ The ball is either black or white.

Some philosophers claim that there is some sense in which we should have $\Pr(X \mid Y) = \frac{1}{2}$ here — as an extra-systematic constraint, of course.

But, intuitively, it’s a different sort of constraint than the one in our last example. In our last example “biased” was itself a probabilistic concept.

Here, there is no probabilistic extra-systematic content involved.

As such, if some extra-systematic probabilistic constraint is called for here, it’s not for purely conceptual reasons. I will call this an epistemic extra-systematic constraint (an instance of the “Principle of Indifference”).

This can be motivated by unpacking $\Pr(X \mid Y)$ as (something like) “the degree of confidence one should have in $X$ — if $Y$ were all one knew.”

We’ll come back to this epistemic understanding of probabilities shortly.
We'll come back to the probabilistic issues soon enough. Let’s back up first, and think more about (extra-systematic) truth (simpliciter).

There are various “Theories” or “Philosophical Explications” of truth. I have posted a nice overview by Haack (and the SEP entry by Glanzberg).

I will separate the philosophical theories of truth into two categories:

- Objective Theories of Truth.
  - Correspondence theories.
- Subjective Theories of Truth.
  - Epistemic theories.
  - Coherence theories.
  - Pragmatic theories.

There are also theories that are neutral on the subjective/objective question. For instance, deflationary theories (like the redundancy theory).

Subjective theories of truth do not involve any sort of correspondence between sentences/propositions/beliefs and some realm of truthmakers.

The epistemic theory of truth, for instance, holds that (Alston):

The truth of a truth bearer consists not in its relation to some “transcendent” state of affairs, but in the epistemic virtues the former displays within our thought, experience, and discourse.

Truth value is a matter of whether, or the extent to which, a belief is justified, warranted, rational, well grounded, or the like.

The coherence theory of truth is a instance of the epistemic theory (where coherence with one’s other beliefs is the salient “epistemic virtue”).

The pragmatic theory of holds that “truth is satisfactory to believe”. Basically, a belief is true if believing it “works” for its believer.

We will adopt an objective/realist stance toward truth in this course. I find it hard to understand the other conceptions of truth. Explain.

According to correspondence theories of truth, \( p \) is true if \( p \) corresponds to some truthmaker \( t_p \) (that is, if there exists a truthmaker \( t_p \) for \( p \)).

There are different views on the bearers of truth-values (sentences, propositions, beliefs) and truthmakers (facts, states of affairs).

Moreover, there are different views about whether truthmakers must exist in some “mind-independent realm”. Realists will require that there is a mind-independent realm of truthmakers. Anti-realists will not.

Just as we can talk about \( p \) being true-on-\( \mathbf{I} \), which is synonymous with \( s_i = p \), we can also talk about \( p \) having probability-on-\( \mathbf{M} \).

And, like truth-on-\( \mathbf{I} \), probability-on-\( \mathbf{M} \) is a logical/formal concept.

That is, once we have specified a probability model \( \mathbf{M} \), this logically determines the probability-on-\( \mathbf{M} \) values of all sentences in \( \mathbf{L} \).

Moreover, just as the truth-on-\( \mathbf{I} \) of sentence \( p \) does not imply anything about \( p \)’s truth (simpliciter), neither does the probability-on-\( \mathbf{M} \) of \( p \) imply anything about \( p \)’s probability (simpliciter) — if there be such a thing.

Finally, just as we have different philosophical “theories” of truth, we will also have different philosophical “theories” of probability.

And, as in the case of truth, there will be objective theories and subjective theories of probability. However, there will be more compelling reasons for “going subjective” in the probability case than in the truth case.

Ultimately, we will be most interested in the assessment of arguments.