LEGAL PROBABILISM

MARCELLO DI BELLO – ASU

HANDOUT #1 – AUGUST 22, 2014

0. WHAT IS LEGAL PROBABILISM?

> "Legal probabilism is a research program that relies on probability theory to analyze, model and improve the evaluation of evidence and the process of decision-making in trial proceedings."
>
> "Legal probabilism can also be understood as a far reaching research program that aims to analyze—by means of probability theory—the trial system as a whole, including institutions such as the jury system and trial procedure."
>
> (Legal Probabilism in the *Stanford Encyclopedia of Philosophy*)

Key idea is to deploy probability theory to:

a. better assess and weigh evidence;

b. guide decision-making in trial proceedings; and

c. analyze and improve the trial system as a whole.

QUESTION: Does this seem a promising research program? Why (yes/no)?

1. THE PHILOSOPHY OF PROBABILITY

What do we mean when we say *there is x percent chance/probability that...*?

> *Classical interpretation*: "The theory of chances consists in reducing all events of the same kind to a certain number of equally possible cases, that is to say, to cases whose existence we are equally uncertain of,[1] and in determining the number of cases favorable to the event whose probability is sought. The ratio of this number to that of all possible cases is the measure of this probability, which is thus only a fraction whose numerator is the number of favorable cases, and whose denominator is the number of all possible cases." (Laplace (1814), *Philosophical Essay of Probabilities*)
>
> *Frequency interpretation:* "We define probability as the limit of a frequency within an infinite sequence." (Reichenbach (1949), *The Theory of Probability*)[2]

---

[1]This reduction rests on the so-called principle of indifference (but see Bertrand's paradox).

[2]Here is another formulation: the probability of an event is the limit of the relative frequency with which the event occurs, assuming repeated trials under the same conditions

Subjective (betting) intepretation: "Let us suppose that an individual is obliged to evaluate the rate p at which he would be ready to exchange the possession of an arbitrary sum S (positive or negative) dependent on the occurrence of a given event E, for the possession of the sum pS; we will say by definition that this number p is the measure of the degree of probability attributed by the individual considered to the event E." (de Finetti (1980), 'Foresight. Its Logical Laws, Its Subjective Sources')

Evidential inteoretation: "Given a scientific hypothesis h, we can intelligibly ask: how probable is h on present evidence? We are asking how much the evidence tells for or against the hypothesis. We are not asking what objective physical chance or frequency of truth h has. A proposed law of nature may be quite improbable on present evidence even though its objective chance of truth is 1. That is quite consistent with the obvious point that the evidence bearing on h may include evidence about objective chances or frequencies. Equally, in asking how probable h is on present evidence, we are not asking about anyone's actual degree of belief in h. Present evidence may tell strongly against h, even though everyone is irrationally certain of h." (Williamson (2000), *Knowledge and Its Limits*)

Model-based interpretation: "Probability is just a property of a mathematical model intended to describe some features of the natural world. For the model to be useful, it must be shown to be in good correspondence with the system it describes. That is where the science comes in." (Stark and Freeman (2016), 'What is the Chance of an Earthquake?')

QUESTION: Which interpretation is the most fitting for trial proceedings?

2. THE MATHEMATICS OF PROBABILITY

P is a probability function provided:

(normality) $0 \leq P(A) \leq 1$, for any proposition A;

(certainty) $P(\top) = 1$, with $\top$ any tautology; and

(additivity) $P(A \vee B) = P(A) + P(B)$, with A, B inconsistent propositions.

The *conditional probability* of A given some other proposition B is defined as follows:

(conditional probability) $P(A|B) = \frac{P(A \wedge B)}{P(B)}$.

Simple corollaries follow, such as:

(negation) $P(\neg A) = 1 - P(A)$;

(product) $P(A \wedge B) = P(A|B)P(B)$;

(product*) $P(A \wedge B) = P(A)P(B)$ if $P(A|B) = P(A)$;

Notation: The symbol '$\neg$' stands for negation, '$\wedge$' for conjunction, and '$\vee$' for disjunction.

Watch out for the inversion fallacy: This is a confusion between $P(A|B)$ and $P(B|A)$.

Example: If you tested positive for a disease and the test returns a positive result in 90% of the cases of disease, that is,

$$P(TestPositive|Disease) = 90\%,$$

this does not imply that

$$P(Disease|TestPositive) = 90\%.$$

To equate the two probabilities is to commit the inversion fallacy.

Prosecutor's Fallacy: When the inversion fallacy is committed in the context of criminal trials, it takes the name of prosecutor's fallacy.

3. COLLINS – MAIN OPINION

Five objections to the introduction of statistical evidence:
(a) statistical estimates were mere guesses;
(b) appeal to independence for application of product rule was unjustified;
(c) the couple involved in the crime might not have had the distinctive features F;
(d) distinctive features F might not be unique;
(e) numbers might have distracted and confused the jury.

Suppose objections (a) and (b) are addressed. Consider this argument from the prosecutor
(step 1) Distinctive features F have a probability of 1 in 12 million in California couples.
(step 2) An innocent couple could have features F with a 1 in 12 million probability.
(step 3) Collins who have features F are innocent with a 1 in 12 million probability.
(step 4) Collins are guilty with a high probability (the complement of 1 in 12 million).
What is wrong with this argument? Which step is problematic?

QUESTION: What proposition, exactly, can be assigned a probability of 1 in 12 million? Is it the probability of innocence? If not, how do we use the figure 1 in 12 million to arrive at the probability of guilt/innocence?

A suggestion is contained in the mathematical appendix in *Collins* (see below).

4. COLLINS – MATHEMATICAL APPENDIX

The Court's calculations rely on the *binomial distribution* formula. To illustrate, let p be the probability of event E, on the supposition that the probability of E does not change as we iterate our experiment (e.g. the probability of getting heads does not change as we keep tossing our coin). The binomial distribution formula allows us to calculate the probability of getting a k number of occurrences of E over n iterations, with event E having a probability of p (e.g. the binomial distribution allows us to calculate the probability of getting a k number of heads over an n number of coin tosses, with the event 'heads' having a probability of one half), as follows:

$$\binom{n}{k} p^k (1-p)^{n-k}.$$

The Court was interested in the probability that there are a k number of couples with features F among California couples (assuming F has probability 1 in 12 million; so objections (a) and (b) above are set aside here). In particular, he Court was interested in the probability that *exactly one couple* among California couples had features F, so this is the case $k = 1$. By putting $k = 1$, the binomial distribution formula becomes:

$$np (1-p)^{n-1}.$$

QUESTION: How did the Court use the binomial distribution formula to arrive at the 0.4 probability? Did the Court use the correct formula here? Is possessing features F among California couples like getting heads on a series of coin tosses? Why not?