VI Probability

- what is it?
- different points of view
- mine, however probability is assigned (and there are different methods) $P(A)$ is a measure of belief that the true value of some unknown is in A.
- whose belief?
- belief given that the basic probability model has been accepted as an adequate reflection of reality $(\Omega, \mathcal{B}, P)$
- how do we know the probability model is correct?
- we don't, but we can use data to check the model
- the best we can do
- there is nothing that can tell us that any probability model is absolutely correct
- but, probability gives a context in which we can formulate the correct rules of reasoning to answer the questions of interest.

① axioms of probability

- sample space Ω, σ-field $\mathcal{F}$, $P : \mathcal{F} \to [0,1]$ s.t.
  (i) $P(\Omega) = 1$
  (ii) if $A_1, A_2, \ldots \in \mathcal{F}$ are mut. disjoint then
  $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$

- another axiom. (first principle of inference)

- principle of conditional probability

- if we are told that $\omega \in C \in \mathcal{F}$ then replace P by $P(\cdot \mid C)$ where
  $P(\cdot \mid C) : \mathcal{F} \to [0,1]$ is defined by
  $P(A \mid C) = P(A \cap C)/P(C)$.
  (assuming $P(C) > 0$)

- note may have $P(A \mid C) < P(A)$ or
  $P(A \mid C) \geq P(A)$

- how is the information $\omega \in C$ generated?

- have to be careful

- it is necessary that there be an information processor $\Xi : \Omega \to \Xi$
  s.t. ξ_0 is observed and $C = \Xi^{-1}\{\xi_0\}$

(2) Principle of Insufficient Reason (Principle of Indifference)

- Laplace

- if $\omega \in \Omega = \{\omega_1, \ldots, \omega_m\}$ and we have no reason to distinguish among the ω_j as possibilities for the true value of ω then assign $P(\{\omega_i\}) = 1/m$

eg The Three Prisoners Problem

- prisoners I, II and III of which two will be executed and one will live

- prisoner I asks the jailer to tell him the name of one of the other prisoners who will be executed as he knows one will be

- the jailer responds II

- prisoner I reasons as follows, initially my prob. of living is 1/3 but now it is 1/2 (principle of indifference) and so he is happy

- is this correct ?

- what is the information processes used by the jailer ?

- we don't know and there are several candidates

- $\Omega = \{(I, II, III), (I, III, II), (II, I, III), (II, III, I),$
  $(III, I, II), (III, II, I)\}$ 1st coord. not executed

- ~~note~~ - we don't really distinguish between permutations of last 2 coordinates

- A = "I is not executed" = $\{(I, II, III), (I, III, II)\}$

- what corresponds to C = "II will be executed"

- two (deterministic) information processors

$$
\begin{array}{lllll}
 & & (1) & & (2) \\
\Xi \; (I, II, III) & = & II & \text{or} & III \\
(I, III, II) & = & II & \text{or} & III \\
(II, I, III) & = & III & & III \\
(II, III, I) & = & III & & III \\
(III, I, II) & = & II & & II \\
(III, II, I) & = & II & & II
\end{array}
$$

$C = \Xi_1^{-1}\{II\} = \{(I, II, III), (I, III, II), (III, I, II), (III, II, I)\}$

$C = \Xi_2^{-1}\{II\} = \{(III, I, II), (III, II, I)\}$

- then with uniform prob. on Ω

$$
P(A|C) = \begin{cases} 1/2 & \text{when using } \Xi_1 \\ 0 & \text{when using } \Xi_2 \end{cases}
$$

- so what prob. is correct?

- suppose jailer randomly picks $\Xi_{①}$ with prob. p

- then $P(A|C) = \frac{p}{1+p} \in [0, 1/2]$
  and we don't know p so we don't know $P(A|C)$

- so there is no justification for I to claim they are more certain that they won't be executed after learning II will be.

- in reality the prob stays the same

- see Monte Hall problem in the text.

3 Subjective Probability

(1) Qualitative probability

- $\Omega = \{\omega_1, \ldots, \omega_m\}$, $\mathcal{F} = 2^{\Omega}$
- a preference ordering is specified on the elements of $\mathcal{F}$ s.t. $A \preceq B$ "A is not more probable than B"
- conditions can be stated for $\preceq$ (see book) s.t. when these are satisfied then $\exists P$ on $\mathcal{F}$ s.t. P agrees with $\preceq$ in the sense that $P(A) \le P(B)$ whenever $A \preceq B$
- but how to specify P?
- suppose $\{\omega_1\} \preceq \{\omega_2\} \preceq \cdots \preceq \{\omega_m\}$ specify $f : \{1, \ldots, m\} \to [0, \infty)$ increasing s.t. $f(i) = 0$ when $\{\omega_i\} \equiv \phi$ and otherwise $f(i+1)/f(i)$ = degree to which ω_{i+1} is more probable than ω_i and then put $P(\{\omega_i\}) = f(i) / \sum_{j=1}^{m} f(j)$

(2) Probability via betting (de Finetti)

- Ω finite, $\omega \in \Omega$ unknown
- $X: \Omega \to \mathbb{R}$ called a gamble where buyer of gamble receives $X(\omega)$ units from seller of gamble
- so if $X(\omega) < 0$ buyer pays $X(\omega)$ to seller
- a better / bookie will buy or sell any finite combination of gambles
- what price will you use to buy/sell a gamble.

Lemma For any (rational) gambler the price ~~at which they will sell~~ X is the price at which they will buy X

Proof: Suppose you will buy X for p_b and sell it for p_s where $p_b > p_s$

$(X(\omega) - p_b) + (p_s - X(\omega)) = p_s - p_b < 0$ $\forall \omega$ which is irrational as you would have a sure loss so $p_b \le p_s$. Now suppose $p_b < p_s$.

Then for any $\varepsilon > 0$ you would not buy X for

$p_b + \varepsilon$ and would not sell it for $p_s - \varepsilon$. But

$(X(\omega) - p_b - \varepsilon) + (p_s - \varepsilon - X(\omega)) = p_s - p_b - 2\varepsilon$

> 0 for ε small enough and this implies the gambler wouldn't accept a sure gain.

Therefore $p_b = p_s$.

- a couple of assumptions
  - (1) everybody has a price at which they will buy or sell any gamble
  - (2) every rational person will avoid a sure loss and accept a sure gain
- let Price(X) = price of X

Def Price is <u>coherent</u> if $\forall$ gambles $X_1, \ldots, X_m, X_{m+1}, \ldots, X_{m+n}$

$$\sup_{\omega \in \Omega} \left\{ \sum_{i=1}^{m} (X_i(\omega) - \text{Price}(X_i)) + \sum_{i=1}^{n} (\text{Price}(X_{m+i}) - X_{m+i}(\omega)) \right\} \geq 0$$

- so Price is coherent iff Price avoids a sure loss on any finite combination of gambles

<u>Theorem</u> Price is coherent iff there is a prob. measure on 2^{Ω} s.t. $\text{Price}(X) = E_p(X)$.

- so for any $A \subseteq \Omega$ we have,

$$P(A) = E_p(I_A) = \text{Price}(I_A)$$

- <u>note</u> - probability arises as a rationality requirement there is no use of repeated performances or randomness

13) <u>Principle of No Arbitrage</u>

- there are no situations when an investor can invest nothing and for which there is a possibility of a gain with no possibility of a loss

- this leads to Black-Scholes formula for pricing options (puts and calls)

- again no notion of repeated performances or randomness

4 Frequentist Probability

- long-run relative frequency

- there is a physical system that can be repeatedly performed and gives values in a set $\mathcal{X}$

- let $x_1, x_2, \ldots$ be a sequence of performances

- define, for $A \subseteq \mathcal{X}$, $P(A) = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} I_A(x_i)$ provided this limit exists.

- what does it mean to say the limit exists?

- but the existence of this limit is not enough (e.g. just turn a coin over and over)

- the system must also behave "randomly"

- we could call systems that satisfy these two criteria <u>random systems</u>

eg games of chance

- we believe that performed correctly are random systems but we can never prove that they are eg - suppose the following sequence is produced of values in {0,1,2,...,9}

0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 1, 0, 1, 1, 1, 2, 1, 3, 1, ...

- we can prove that all frequencies will converge and in fact this sequence satisfies all the tests we could think of to test it as an outcome from iid $f(i) = 1/10$.

- Champernowne's sequence

- what is randomness?

Kolmogorov (1963): a sequence is random when you can't find a short program to generate it

- so there are many ways to think about probability
- but P(A) is always measuring the belief that the value of some unknown is in A
- we want a theory of inference that doesn't depend on any particular way of probabilities being assigned.