(I) Long Run Frequency / Limiting Frequency Definition

$$P(A) := \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}_{\omega_i \in A}$$

(II) Karl Popper, 1957

Objects have 'inherent' dispositions towards outcomes

"propensity"

Propensity induces the long run freq.

Radioactive U238

$$\underbrace{P(\text{U238 atom explodes} < 4.5 \text{ Billion year})}_{\text{can be calculated explicitly if you understand quantum mechanics}} = \frac{1}{2}$$

Problems

(1) For most random experiments, we don't know how to calculate the propensities of ω's

(2) Not general
$P(\text{OJ Simpson guilty})$

I, II are objectivist theories

(III) Subjectivist Definition: Everyone uses their own evidence, biases, intuition to come up with their own estimate of uncertainty

$P_{Adam}(H) = 0.5$ Ramsey, 1926

$P(\text{Newton's } F = ma \text{ is true})$ de Finetti, 1928

Problem

No one answer

Conclusion

No accepted definition of probability

What is randomness? → Before 1920's: "randomness is merely ignorance"

→ After 1920's: the Universe is random (double slit experiment)

Kolmogorov 1930's = Mathematical Definition of Probability

Assume $\exists\ \Omega \neq \emptyset$ P is a set function satisfying the 3 conditions:

Axioms

(A) $P(\Omega) = 1$

(B) $\forall A \subseteq \Omega \quad P(A) \geq 0$

(C) If $A_1, A_2, \ldots$ are disjoint $\Rightarrow P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$

Proofs - using the Axioms

Thm I Complement Rule

$P(A) = 1 - P(A^c)$

$\Omega = A \cup A^c$ and $\{A, A^c\}$ are disjoint

$P(\Omega) = P(A \cup A^c)$

$P(\Omega) = P(A) + P(A^c)$ by Ⓒ

$1 = P(A) + P(A^c)$ by Ⓐ

$P(A) = 1 - P(A^c)$ ✓

Thm II

$P(\emptyset) = 0$

$P(\emptyset) = 1 - P(\emptyset^c)$ by thm I

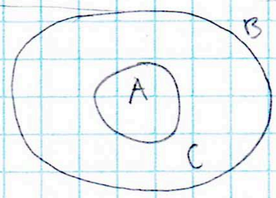
$= 1 - P(\Omega)$ set theory

$= 1 - 1$ by Ⓐ

$= 0$ ✓

Thm III

$A \subseteq B \Rightarrow P(A) \subseteq P(B)$



$C = B \setminus A$

$C \cap A = \emptyset$

$P(B) = P(A \cup C) = P(A) + P(C)$

$P(B) - P(A) = P(C) \geq 0$ by Ⓑ

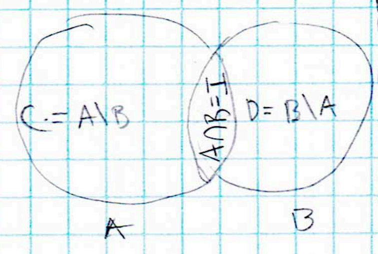
$P(B) - P(A) \geq 0$

$P(B) \geq P(A)$

$P(A) \leq P(B)$ ✓

Thm 4
Law of Inclusion-Exclusion

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$



$P(C \cup I) = P(C) + P(I) = P(A)$

$P(D \cup I) = P(D) + P(I) = P(B)$

by Ⓒ

$P(A \cup B) = P(C \cup I \cup D) = P(C) + P(I) + P(D)$ by Ⓒ

$= (P(A) - P(I)) + \cancel{P(I)} + (P(B) - \cancel{P(I)})$

$= P(A) + P(B) - P(I)$

$= P(A) + P(B) - P(A \cap B)$ ✓

Thm 5

$|\Omega| < \infty$ if $P(\omega_i) = \frac{1}{|\Omega|}\ \forall \omega_i \Rightarrow P(A) = \frac{|A|}{|\Omega|}$

let $n = |A| < \infty$

since $A \subseteq \Omega \Rightarrow |A| \leq |\Omega|$

$A = \{\omega_1, \omega_2, \ldots, \omega_n\}$

$A = \{\omega_1\} \cup \{\omega_2\} \cup \ldots \cup \{\omega_n\}$

$P(A) = P\left(\bigcup_{i=1}^{n} \{\omega_i\}\right) = \sum_{i=1}^{n} P(\{\omega_i\})$ by Ⓒ

$= \sum_{i=1}^{n} \frac{1}{|\Omega|} = \frac{n}{|\Omega|} = \frac{|A|}{|\Omega|}$

Conditional Probability

$n = 1000$ people

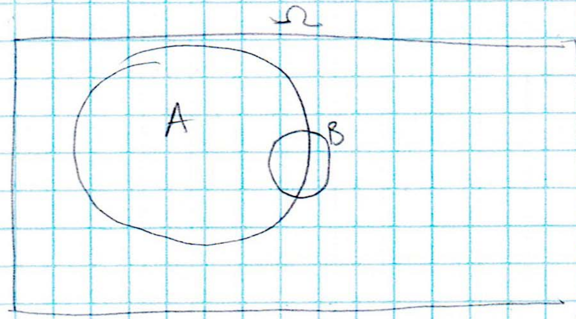
200 smokers (A = smoking)

60 lung cancer (B = lung cancer)

36 smoke and gets lung cancer ($A \cap B$)

Assume for illustrative purpose

$$\left.\begin{array}{l} P(A) = 0.2 \\ P(B) = 0.06 \\ P(A \cap B) = 0.036 \end{array}\right\}$$



What is the probability of lung cancer among smokers

new universe

pipe denotes "conditioning" on event A

cond. prob. $P(A \mid B)$

$\Omega' = A$

"new universe"