1 Introduction to Probability Theory

P.T. provides

- concepts to speak about uncertainty (games of chance, recurring events with some patterns of variation, e.g., measurements, balls in Galton board, etc.)
- methods to quantify uncertainty

Example: Rolling a die D

What is $P(D=6)$? $P(D \geq 4)$? $P(D \leq 4)$?

Meaning:

1) In the long run, $1/6$ of throws result in a 6
2) There is a chance of 1 in 6 that <u>this throw</u> results in a 6

1) = <u>frequentist</u> view 2) = subjective (Bayesian) view

No consequence for mathematical theory

1.1 Events

Experiments:

$\mathcal{S}$ sample space (= set of possible outcomes)

$\mathcal{S}$ is known, but not specific outcome

Rolling a die: $\mathcal{S} = \{1, 2, \dots, 6\}$

Rolling two dice: $\mathcal{S} = \{(1,1), (1,2), \dots, (1,6), (2,1), \dots\dots, (6,6)\}$

Galton board: $\mathcal{S} = \{1, 2, \dots, n, n+1\}$

$\mathcal{S}$ can be infinite!

Events:

throw an even number $\mathcal{E}_{even} = \{2, 4, 6\}$

throw two equal numbers $\mathcal{E}_{equal} = \{(1,1), \dots, (6,6)\}$

In general:

$$\mathcal{E} \subseteq \mathcal{S}$$

$\mathcal{E}$ occurs if outcome $\in \mathcal{E}$

Set operations on events

$\mathcal{E} \cup \mathcal{F}$ union ("or")

$\mathcal{E} \cap \mathcal{F}$, $\mathcal{E}\mathcal{F}$ intersection ("and")

$\overline{\mathcal{E}}$ complement ("not")

Example:

$$\mathcal{E}_{prime} := \{2,3,5\}$$

$$\mathcal{E}_{even} \cup \mathcal{E}_{prime} = \{2,3,4,5,6\}$$

$$\mathcal{E}_{even} \cap \mathcal{E}_{prime} = \{2\}$$

$$\overline{\mathcal{E}_{prime}} = \{1,4,6\}$$

Special case: $\emptyset$ <u>empty event</u>

Terminology and examples:

$\mathcal{E} \cap \mathcal{F} = \emptyset$ $\quad$ $\mathcal{E}$ and $\mathcal{F}$ are <u>mutually exclusive</u>

$\overline{\mathcal{S}} = \emptyset, \quad \overline{\emptyset} = \mathcal{S}$

$\overline{\overline{\mathcal{E}}} = \mathcal{E}$

$\overline{\mathcal{E} \cup \mathcal{F}} = \overline{\mathcal{E}} \cap \overline{\mathcal{F}}, \quad \overline{\mathcal{E} \cap \mathcal{F}} = \overline{\mathcal{E}} \cup \overline{\mathcal{F}}$ (De Morgan Rules)

$\mathcal{E} \subseteq \mathcal{F}$ $\quad$ $\mathcal{E}$ is <u>contained</u> in $\mathcal{F}$

$\mathcal{E} \equiv \mathcal{F} \Leftrightarrow \mathcal{E} \subseteq \mathcal{F}$ and $\mathcal{F} \subseteq \mathcal{E}$

$\mathcal{E}$ and $\mathcal{F}$ are <u>equivalent</u>

1.2. Axioms of Probability

$P(\mathcal{E})$ probability of $\mathcal{E}$, is a real number

A1: $0 \leq P(\mathcal{E}) \leq 1$

A2: $P(\mathcal{S}) = 1$

A3: If $\mathcal{E}_1, \mathcal{E}_2, \dots$ are mutually exclusive (i.e., $\mathcal{E}_i \cap \mathcal{E}_j = \emptyset$ for $i \neq j$),

then $P(\mathcal{E}_1 \cup \dots \cup \mathcal{E}_n) = P(\mathcal{E}_1) + \dots + P(\mathcal{E}_n)$

$$= \sum_{i=1}^{n} P(\mathcal{E}_i) \qquad \text{f.a. } n$$

(also for infinite sums)

What is $P(\mathcal{E}_{even} \cup \{1\})$?

What is $P(\emptyset)$? Why?

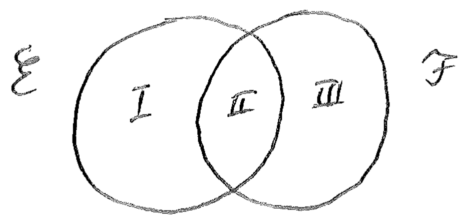
Proposition 1: $P(\mathcal{E}) = 1 - P(\bar{\mathcal{E}})$

Proof: $\mathcal{E} \cap \bar{\mathcal{E}} = \emptyset$, $\mathcal{E} \cup \bar{\mathcal{E}} = \mathcal{S}$

$\Rightarrow P(\mathcal{E}) + P(\bar{\mathcal{E}}) = P(\mathcal{E} \cup \bar{\mathcal{E}}) = P(\mathcal{S}) = 1$

$\Rightarrow P(\mathcal{E}) = 1 - P(\bar{\mathcal{E}})$ $\square$

What about $P(\mathcal{E} \cup \mathcal{F})$ in general?



$$P(E \cup F) = P(\mathrm{I}) + P(\mathrm{II}) + P(\mathrm{III})$$

$$P(E) = P(\mathrm{I}) + P(\mathrm{II})$$

$$P(F) = P(\mathrm{II}) + P(\mathrm{III})$$

$\Rightarrow$ $P(E \cup F) = P(E) + P(F) - P(\mathrm{II})$,

where $P(\mathrm{II}) = P(E \cap F)$

Proposition 2: $P(E \cup F) = P(E) + P(F) - P(E \cap F)$

Example 3. 68% of adults drink beer, 49% wine, 35% both, wine and beer. What percentage drink neither?

$$P(B \cup W) = P(B) + P(W) - P(BW)$$

$$= 0.68 + 0.49 - 0.35 = 1.17 - 0.35$$

$$= 0.82$$

$\Rightarrow$ $P(\overline{B \cup W}) = 0.18$

Definition 4. The Odds of $\mathcal{E}$ is

$$\frac{P(\mathcal{E})}{P(\bar{\mathcal{E}})} = \frac{P(\mathcal{E})}{1 - P(\mathcal{E})}$$

Says how much more likely $\mathcal{E}$ is than $\bar{\mathcal{E}}$.

Example: Odds of rolling a 2 with a die

$$P(\mathcal{E}) = \frac{1}{6} \Rightarrow \frac{P(\mathcal{E})}{1 - P(\mathcal{E})} = \frac{\frac{1}{6}}{1 - \frac{1}{6}} = \frac{\frac{1}{6}}{\frac{5}{6}}$$

$$= \frac{1}{5}$$

1.3 Uniformity.

In many cases all outcomes are equally likely.
(Only possible if $\# S < \infty$.)

$\# S = n$, say $S = \{1, \dots, n\}$

$\Rightarrow \quad P(\{1\}) = P(\{2\}) = \dots = P(\{n\}) = p$

$\overset{A1, A2}{\Rightarrow} \quad 1 = P(S) = P(\{1\}) + \dots + P(\{n\}) = n p$

$\Rightarrow \quad p = \frac{1}{n} = P(\{i\}), \quad 1 \le i \le n$

Generalize: $E \subseteq S \Rightarrow P(E) = \dfrac{\# E}{n}$

Counting Principle:

\# outcomes : throw 2 dice, 3 dice

or: throw a die, pick a card off a stack of 32

Generalize: Consider Experiment 1 followed by Experiment 2

E1 with m outcomes, E2 with n outcomes

\# outcomes of "E1 then E2" is $m \cdot n$

Idea: outcomes form matrix

$$\begin{matrix} (1,1) & (1,2) & \cdots & (1,n) \\ (2,1) & (2,2) & \cdots & \vdots \\ \vdots & \vdots & & \vdots \\ (m,1) & (m,2) & & (m,n) \end{matrix}$$

Generalize: r experiment with $n_1, \dots, n_r$ outcomes

$\Rightarrow n_1 \cdot \dots \cdot n_r$ outcomes of combination