Definition of Probability

Ruben Zamar
Department of Statistics
UBC

January 8, 2019

RANDOM EXPERIMENT

- The outcome cannot be determined beforehand
- Examples
  - number of shots needed to decide a tennis point
  - yield of a chemical process
  - max-wind speed in Vancouver in 2015
  - your final grade in STAT 302

SAMPLE SPACE

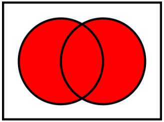
- List of all the possible outcomes of a random experiment
- Denoted by Ω
- Examples
    - number of shots: $\Omega = \{1, 2, 3, 4, ...\}$
    - yield of a chemical process: $\Omega = [0, 100]$ in percentage
    - max-wind speed: $\Omega = [0, \infty)$ [or $[0, 1000)$ in km/hour]
    - final grade: $\Omega = [0, 100]$

EVENT

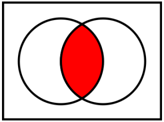
- Subsets of Ω are called events
- Denoted by A, B, C,etc.
- A occurs if $\omega \in A$
- Examples
  - $A = \{\text{at most 3 shots}\} = \{1, 2, 3\}$
  - $B = \{\text{between 20 and 40 percent}\} = [20, 40]$
  - $C = \{\text{over 100 km/hour}\} = [100, \infty)$
  - $D = \{\text{an ace}\} = [80, 100]$

Boolean algebra (set operations)

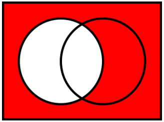
Union: $A \cup B$,



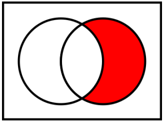
Intersection: $A \cap B$,



Complement: A^c,



Difference: $B \backslash A = B \cap A^c$



De Morgan's Laws

$$(A \cup B)^c = A^c \cap B^c$$
$$(A \cap B)^c = A^c \cup B^c$$

Commutative, Associative and Distributive Properties

$$A \cup B = B \cup A \quad \text{(commutative)}$$

$$A \cup B \cup C = (A \cup B) \cup C = A \cup (B \cup C) \quad \text{(associative)}$$

$$A \cap B \cap C = (A \cap B) \cap C = A \cap (B \cap C) \quad \text{(associative)}$$

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C) \quad \text{(distributive)}$$

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C) \quad \text{(distributive)}$$

Domain for the Probability Function

- Denoted by $\mathcal{F}$
- A collection of events

- Required properties for $\mathcal{F}$

  - ϕ and $\Omega \in \mathcal{F}$
  - $A \in \mathcal{F} \implies A^c \in \mathcal{F}$
  - $A_n \in \mathcal{F} \implies \cup_{n=1}^{\infty} A_n \in \mathcal{F}$

- Examples
  - $\mathcal{F} = \{\phi, \Omega\}$
  - $\mathcal{F} = \{A, A^c, \phi, \Omega\}$
  - all subsets of Ω : $\mathcal{F} = 2^{\Omega}$ (this will be used in this course)

DISCUSSION

- $A_n \in \mathcal{F} \implies \cap_{n=1}^{\infty} A_n \in \mathcal{F}$ (show this)
- If $\mathcal{F}_\alpha$are prob domains then $\mathcal{F} = \cap_\alpha \mathcal{F}_\alpha$ is also a prob domain (why?)
- $\mathcal{C} \subset 2^\Omega$, $\mathcal{F}(\mathcal{C}) = \cap_{\mathcal{C} \subset \mathcal{F}_\alpha} \mathcal{F}_\alpha$
- Borel sets $= \mathcal{F}$ (open intervals)

PROBABILITY FUNCTION

$P : 2^{\Omega} \rightarrow [0, 1]$ satisfying:

1. $P(\Omega) = 1$
2. $P(A) \geq 0$, for all $A \in \mathcal{F}$
3. A_n disjoint $\implies P\left(\cup_{n=1}^{\infty} A_n\right) = \Sigma_{n=1}^{\infty} P(A_n)$

Probability space: $(\Omega,\ P)$

SIMPLE RESULT

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

Proof:

$$A^c \cup A = \Omega$$

$$\Longrightarrow \qquad P(A^c) + P(A) = P(\Omega) = 1$$

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

SIMPLE RESULT

$$A \subset B \implies P(A) \leq P(B)$$

Proof:

$$A \cup A^c = \Omega$$

$$B \cap (A \cup A^c) = B \cap \Omega = B$$

$$(B \cap A) \cup (B \cap A^c) = A \cup (B \cap A^c) = B$$

$$P[A \cup (B \cap A^c)] = P(A) + P(B \cap A^c) = P(B)$$

$$\implies \quad P(A) \leq P(B)$$

SIMPLE RESULT

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

Proof:

$$A \cup B = A \cup (B \cap A^c)$$

$$\Longrightarrow P(A \cup B) = P(A) + P(B \cap A^c) \qquad (1)$$

On the other hand,

$$P(B) = P(B \cap A^c) + P(B \cap A)$$

$$\Longrightarrow P(B \cap A^c) = P(B) - P(B \cap A) \qquad (2)$$

From (1) and (2)

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

INCLUSION-EXCLUSION FORMULA

The formula can be extended for the union of $m \geq 2$ sets. For instance,

$$
\begin{aligned}
P(A_1 \cup A_2 \cup A_3) = {} & P(A_1) + P(A_2) + P(A_3) \quad \text{inclusion} \\
& - P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) \quad \text{exclusion} \\
& + P(A_1 \cap A_2 \cap A_3) \quad \text{inclusion}
\end{aligned}
$$

INCLUSION-EXCLUSION FORMULA

More generally,

$$
\begin{aligned}
P\left(\cup_{i=1}^{n} A_i\right) = & \sum_{1 \le i \le n} P\left(A_i\right) \quad \text{inclusion} \\
& - \sum_{1 \le i < j \le n} P\left(A_i \cap A_j\right) \quad \text{exclusion} \\
& + \sum_{1 \le i < j < k \le n} P\left(A_i \cap A_j \cap A_k\right) \quad \text{inclusion} \\
& - \sum_{1 \le i < j < k < l \le n} P\left(A_i \cap A_j \cap A_k \cap A_l\right) \quad \text{exclusion} \\
& \vdots \\
& + (-1)^{n-1} P\left(A_1 \cap A_2 \cdots \cap A_n\right)
\end{aligned}
$$

INCLUSION-EXCLUSION FORMULA

More useful when there is "symmetry":

$$P(A_1 \cap A_2 \cap \cdots \cap A_k) = P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = p_k$$

for all

$$1 \leq i_1 < i_2 < \cdots < i_k \leq n$$

In this case:

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} p_k$$

BOOLE's INEQUALITY

- $$P\left(\cup_{i=1}^{n} A_i\right) \;\leq\; \sum_{i=1}^{n} P\left(A_i\right)$$

- If the events are disjoint, then the inequality becomes an equality.

- Easy to prove for the case of $n = 2$

- For general n use induction

EQUALLY LIKELY OUTCOMES

$$P(A) = \frac{\text{number of outcomes in } A}{\text{number of outcomes in } \Omega} = \frac{\#\ A}{\#\ \Omega}$$

$$= \frac{\text{“favorable” outcomes}}{\text{possibles outcomes}}$$

Combinatorial theory ...