MATH 302  Introduction to Probability

Yinan Spinka
yinan@math.ubc.ca
www.math.ubc.ca/~yinan/math302

Office hours:  
M 15-16
W 16-17
F 15-16  LSK 300C

Grading:
HW 9 out of 10  20%
Mid-term         30%
Final exam       50%
Book: "Introduction to Probability"
by Anderson, Seppäläinen, Valkó

Examples of real/conceptual experiments whose understanding involves prob. theory:

1) Toss 2 dice
2) Deal a poker hand
3) Spin a roulette wheel
4) Lifetime of a radioactive isotope
5) Position of a particle undergoing diffusion
6) # of accidents per year at a busy intersection.
7) 130 million votes are cast in an election. "Determine" the winner before the evening news.
In this course, we will:

1. Learn how to formalize randomness in the language of math.
2. Prove/investigate some important features of random systems.
3. Get to know some widely used models for random systems.

Example: Toss 2 "fair" dice. What is the probability that the total is 7?

Sol: Let's write out all possible outcomes. Suppose one die is red and one blue. We write first the value of the red die and then the blue.
red is 1 → (1,1) (1,2) (1,3) (1,4) (1,5) (1,6)
red is 2 → (2,1) (2,2) (2,3) (2,4) (2,5) (2,6)
red is 6 → (6,1) (6,2) (6,3) (6,4) (6,5) (6,6)

blue is 1

blue is 2

blue is 6

36 possible outcomes
6 outcomes have total 7

\[ \frac{6}{36} = \frac{1}{6} \]
**Terminology**: The set of all possible outcomes is called the **sample space**. We usually denote it by $S$.

In the example, $S = \{(a,b) : a, b \in \{1, 2, \ldots, 6\}\}$

The phrase "the total is 7" serves to single-out some outcomes in $S$.

**Terminology**: A subset $E$ of $S$ is called an **event**.

In the example, "the total is 7" = $E = \{(1,6), (2,5), \ldots, (6,1)\}$
More examples:

(*) Toss a coin 3 times.

Sample space $S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$

$|S| = 8$

(*) $E = \text{"the event that the third toss is tails"} = \{HHT, HTT, THT, TTT\}$

(*) $F = \text{"the event that there is exactly one \textit{head}s"} = \{HTT, THT, TTH\}$
(4) \( G = \{ \text{the event that the third toss is tails and there is exactly one heads} \} = \{ \text{HHT, THT} \} \)

(4) \( H = \{ \text{the event that the third toss is tails or there is exactly one heads} \} = \{ \dot{\ldots}, 3 \} \)

(4) Somebody else tosses a coin 3 times and tells you the number of heads:

sample space \( S = \{0, 1, 2, 3\} \)

(4) Observe the lifetime of a light bulb (in days):

sample space \( S = [0, \infty) \cup \{0\} = [0, \infty] \)

\( E = \{ \text{burns out within a year} \} = [0, 365) \).
\([0, \infty) \neq [0, \infty]\)

"Operations" of events:

1. \(E \subset F\): if \(E\) occurs, then so does \(F\).

2. \(E \cap F\): both \(E\) and \(F\) occur.

3. \(E \cup F\): either \(E\) or \(F\) occurs.

\(S = (0, \infty)\)
\(S = [0, \infty]\)
Last time: sample space $S$
events $E$, subset of $S$

(4) $E \subseteq F$: if $E$ occurs, then $F$ occurs

(4) $E \cap F$: both $E$ and $F$ occur

(4) $E \cup F$: either $E$ or $F$ occurs

(4) $E \cap F = \emptyset$: $E$ and $F$ cannot simultaneously occur

(4) $E^c$: $E$ does not occur

Today: "arithmetical laws" of events

define probability - we would to assign probabilities $P(E)$ to all events $E$. 
"Arithmetic laws" of events:

(i) \((E \cup F) \cap G = (E \cap G) \cup (F \cap G)\)

(ii) \((E \cup F) \cap G = (E \cap G) \cup (F \cap G)\)

(iii) \((E^c)^c = E\), \(S^c = \emptyset\), \(\emptyset^c = S\)

(iv) De Morgan's laws:

\((E \cap F)^c = E^c \cup F^c\)

\((E \cup F)^c = E^c \cap F^c\)

\((\bigcap_{i=1}^n E_i)^c = \bigcup_{i=1}^n E_i^c\)
We would like to assign a number $P(E)$ to each event $E$ in a consistent manner.

**Example**: a fair coin toss

$S = \{ H, T \}$

events: $\{H, T, \emptyset, \{H, T\}\}$

$P(\{H\}) = \frac{1}{2} = P(\{T\})$

$P(\emptyset) = 0$, $P(\{H, T\}) = 1$

**Definition**: a probability (probability measure or distribution) is a way of assigning numbers $P(E)$ to each event $E$ in such a way that the following "axioms" hold:
(I) For any event \( E \), \( 0 \leq P(E) \leq 1 \).

(II) \( P(S) = 1 \)

(III) For any sequence \( E_1, E_2, \ldots \) of pairwise disjoint events, \( E_i \cap E_j = \emptyset \) for all \( i \neq j \)

\[
P(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} P(E_n)
\]

Important example: uniform probability

Suppose \( S \) is finite.

Set \( P(E) = \frac{|E|}{|S|} \) for any \( E \). \( |E| = \#E \) = number of elements in \( E \)

Let's check that the axioms are satisfied.
Axiom (I): for any \( E \), \( 0 \leq |E| \leq |S| \).

Therefore, \( 0 \leq P(E) = \frac{|E|}{|S|} \leq 1 \).

Axiom (II): \( P(\emptyset) = \frac{|\emptyset|}{|S|} = 0 \).

Axiom (III): Suppose \( E_1, E_2, \ldots \) are pairwise disjoint events.

\[
P(\bigcup_{n=1}^{\infty} E_n) = \frac{\bigcup_{n=1}^{\infty} |E_n|}{|S|} = \frac{\sum_{n=1}^{\infty} |E_n|}{|S|} = \frac{n}{|S|} \]

Particular example: fair die roll

\( S = \{1, 2, \ldots, 6\} \)

\( P(E) = \frac{|E|}{6} \)

\( P(1 \text{ or } 6) = \frac{3}{6} = \frac{1}{2} \)
Remarks about definition:

(1) One interpretation of probability is that if we conduct the same experiment again and again, the percentage of times that \( E \) occurs is "close" to \( P(E) \).

(2) When \( S \) is finite, say \( S=\{1,2,...,n\} \), the numbers \( p_i = P(\{i\}) \), \( i=1,...,n \), are sufficient to compute \( P(E) \) for any \( E \). Indeed:

\[
    P(E) = \sum_{i \in E} p_i
\]

Explanation: \( E = \bigcup_{i \in E} \{i\} \) and use axiom (III).
(x) When $S$ is infinite (uncountable), complications may arise:

1. it may happen that $\mathbb{P}(\{s\}) = 0$ for all $s \in S$, but still $\mathbb{P}(E) > 0$ for some $E$.
2. there may be subsets $E \subset S$ that do not have a well-defined prob.