ECE 3530/CS 3130: Engineering Probability & Stats
Lecture Notes Spring 2015

Lecture #2

1 Events as Sets

All probability is defined on sets. In probability, we call these sets *events*. A set is a collection of elements. In probability, we call these *outcomes*. The words we use for events is slightly different than what you did when you learned sets in a math class, see Table 1.

Def'n: *Event*
A collection of outcomes. Order doesn't matter, and there are no duplicates.

Set Theory	**Probability Theory**	**Probability Symbol**
universe	sample space (certain event)	S
element	outcome (sample point)	s
set	event	E
disjoint sets	disjoint events	$E_1 \cap E_2 = \emptyset$
null set	null event	$\emptyset$

Table 1: Set Terminology vs. Probability Terminology

1.1 Introduction

There are different ways to define an event (set):

- List them: $A = \{0, 5, 10, 15, \ldots\}$; $B = \{Tails, Heads\}$.
- As an interval: $[0, 1]$, $[0, 1)$, $(0, 1)$, $(a, b]$. Be careful: the notation overlaps with that for coordinates!
- An existing event set name: $\mathbb{N}$, $\mathbb{R}^2$, $\mathbb{R}^n$.
- By rule: $C = \{x \in \mathbb{R} | x \geq 0\}$, $D = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 < R^2\}$.

1.2 Venn Diagrams

Venn diagrams can be used to pictorially show whether or not there is overlap between two or more sets. They are a good tool for helping remember some of the laws of probability. We don't use them in proofs, however, they're particularly good to develop intuition.

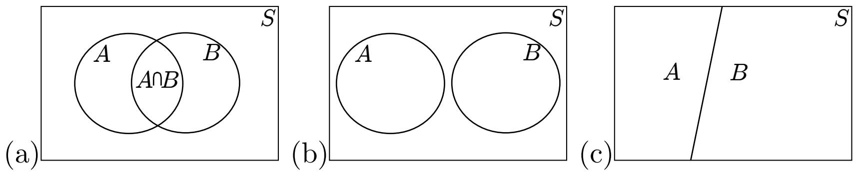

Figure 1: Venn diagrams of sample space S and two sets A and B which (a) are not disjoint, (b) are disjoint, and (c) form a partition of S.

1.3 Important Events

Here's an important event: $\emptyset = \{\}$, the *null event* or the *empty set*.

Here's the opposite: S is used to represent the set of everything possible in a given context, the *sample space*.

- $S = B$ above for the flip of a coin.
- $S = \{1, 2, 3, 4, 5, 6\}$ for the roll of a (6-sided) die.
- $S = \{Adenine, Cytosine, Guanine, Thymine\}$ for the nucleotide found at a particular place in a strand of DNA.
- $S = C$, *i.e.*, non-negative real numbers, for your driving speed (maybe when the cop pulls you over).

1.4 Operating on Events

We can operate on one or more events:

- Complement: $A^c = \{x \in S | x \notin A\}$. We must know the sample space S!
- Union: $A \cup B = \{x | x \in A \,\text{or}\, x \in B\}$. Merges two events together.
- Intersection: $A \cap B = \{x | x \in A \,\text{and}\, x \in B\}$. Limits to outcomes common to both events.

DO NOT use addition to represent the union, and DO NOT use multiplication to represent the intersection. An example of why this is confusing:

$$\{1\} + \{1\} = \{1\}.$$

This leads to one of the most common written mistakes – exchanging unions and plusses when calculating probabilities. Don't write $P[A] + P[B]$ when you really mean $P[A \cup B]$. Don't add sets and numbers: for example, if A and B are sets, don't write $P[A] + B$.

1.5 Disjoint Sets

Def'n: *Disjoint*
Two events A_1 and A_2 are disjoint if $A_1 \cap A_2 = \emptyset$.

Def'n: *Mutually exclusive*
Multiple events $A_1, A_2, A_3, \ldots$ are mutually exclusive if every pair of events is disjoint.

Some disjoint events: $\{1, 2\}$ and $\{6\}$; A and A^c; A and $\emptyset$.

2 Probability

You're familiar with functions, like $f(x) = x^2$, which assign a number output to each number input. Probability assigns a number output to each event input.

2.1 How to Assign Probabilities to Events

As long as we follow three intuitive rules (*axioms*) our assignment can be called a *probability model*.

Axiom 1: For any event A, $P[A] \geq 0$.

Axiom 2: $P[S] = 1$.

Axiom 3: For any two disjoint events A_1 and A_2,

$$P[A_1 \cup A_2] = P[A_1] + P[A_2].$$

The final axiom in the Walpole book is more complicated:

Axiom 3: If $A_1, A_2, A_3, ...$ is a sequence of mutually exclusive events, then

$$P[A_1 \cup A_2 \cup A_3 \cup \cdots] = P[A_1] + P[A_2] + P[A_3] + \cdots$$

However, one may prove the more complicated Walpole Axiom 3 from the three axioms given above. For details, see [Billingsley 1986].

Example: DNA Measurement

Consider the DNA experiment above. We measure from a strand of DNA its first nucleotide. Let's assume that each nucleotide is equally likely. Using axiom 3,

$$P[\{a, c, g, t\}] = P[\{a\}] + P[\{c\}] + P[\{g\}] + P[\{t\}]$$

But since $P[\{a, c, g, t\}] = P[S]$, by Axiom 2, the LHS is equal to 1. Also, we have assumed that each nucleotide is equally likely, so

$$1 = 4P[\{a\}]$$

So $P[\{a\}] = 1/4$.

<u>Def'n:</u> *Discrete Uniform Probability Law*
In general, for event A in a discrete sample space S composed of equally likely outcomes,

$$P[A] = \frac{|A|}{|S|}$$

2.2 Properties of Probability Models

1. $P[A^c] = 1 - P[A]$. Proof:
   First, note that $A \cup A^c = S$ from above. Thus

   $$P[A \cup A^c] = P[S]$$

   Since $A \cap A^c = \emptyset$ from above, these two events are disjoint.

   $$P[A] + P[A^c] = P[S]$$

   Finally from Axiom 2,

   $$P[A] + P[A^c] = 1$$

   And we have proven what was given.

   Note that this implies that $P[S^c] = 1 - P[S]$, and from axiom 2, $P[\emptyset] = 1 - 1 = 0$.

2. For any events E and F (not necessarily disjoint),

   $$P[E \cup F] = P[E] + P[F] - P[E \cap F]$$

   Essentially, by adding $P[E] + P[F]$ we double-count the area of overlap. Look at the Venn diagram in Figure 1(a). The $-P[E \cap F]$ term corrects for this. Proof: Do on your own using these four steps:

   (a) Show $P[A] = P[A \cap B] + P[A \cap B^c]$.
   (b) Same thing but exchange A and B.
   (c) Show $P[A \cup B] = P[A \cap B] + P[A \cap B^c] + P[A^c \cap B]$.
   (d) Combine and cancel.

3. Subset rule: If $A \subset B$, then $P[A] \leq P[B]$. Proof:
   Let $B = (A \cap B) \cup (A^c \cap B)$. These two events are disjoint since $A \cap B \cap A^c \cap B = \emptyset \cap B = \emptyset$. Thus:

   $$P[B] = P[A \cap B] + P[A^c \cap B] = P[A] + P[A^c \cap B] \geq P[A]$$

   Note $P[A \cap B] = P[A]$ since $A \subset B$, and the inequality in the final step is due to the Axiom 1.

4. Union bound: $P[A \cup B] \leq P[A] + P[B]$. Proof: directly from 2 and Axiom 1.

5. Conjunction bound: $P[A \cap B] \leq P[A]$. Proof: $A \cap B$ is a subset of A. This then follows from the subset rule above. See the "Conjunction Fallacy" from lecture 1.