1.3

Wednesday, September 1, 2021 3:22 PM

SECTION 1.3: RANDOM EXPERIMENTS AND PROBABILITIES

1.3.1 Random Experiments

Connecting the notion of probability to set theory:

> Outcome: A result of a random experiment.
> Sample Space: The set of all possible outcomes.
> Event: A subset of the sample space.

or event = collection of outcomes

Union and Intersection:

If A and B are events, then $A \cup B$ and $A \cap B$ are also events.

We observe that $A \cup B$ occurs if A or B occur. (or both)

Similarly, $A \cap B$ occurs if both A and B occur.

Similarly, if $A1, A2, \cdots, An$ are events, then the event $A1 \cup A2 \cup A3 \cdots \cup An$ (union) occurs if at least one of $A1, A2, \cdots, An$ occurs.

The event $A1 \cap A2 \cap A3 \cdots \cap An$ occurs if all of $A1, A2, \cdots, An$ occur.

It can be helpful to remember that the key words "or" and "at least" correspond to unions and the key words "and" and "all of" correspond to intersections.

What is $\underbrace{(A_1 \cup A_2 \cup \cdot \cup A_n)}_{\text{at least one of } A_1, \dots, A_n}^c = \underbrace{A_1^c \cap A_2^c \cap \cdot \cap A_n^c}_{\text{none of } A_1, \dots, A_n \text{ occurs}}$

De Morgan's law

1.3.2 Probability

Now that we have discussed operations on sets (or "events") we can now in a rational way mathematically define "probability" (or sometimes also called "probabiity mesure".

A probability is a measure of plausibility of random events. ($P(A)$ ....etc.). Probability should always lie between 0 and 1.

Here are the rules of probability that ensure that we always obtain sensible results when we apply these rules. These rules apply equally in countable sample spaces as well as in uncountable sample spaces.

> Axioms of Probability:
>
> - Axiom 1: For any event A, $P(A) \geq 0$.
> - Axiom 2: Probability of the sample space S is $P(S) = 1$.
> - Axiom 3: If $A_1, A_2, A_3, \cdots$ are disjoint events, then $P(A_1 \cup A_2 \cup A_3 \cdots) = P(A_1) + P(A_2) + P(A_3) + \cdots$

also write $Pr(A)$

disjoint = mutually exclusive

- Axiom 3: If $A_1, A_2, A_3, \cdots$ are disjoint events, then $P(A_1 \cup A_2 \cup A_3 \cdots) = P(A_1) + P(A_2) + P(A_3) + \cdots$

disjoint = mutually exclusive
≡ no intersections

In summary, if $A1$ and $A2$ are disjoint events, then $P(A1 \cup A2) = P(A1) + P(A2)$. The same argument is true when you have n disjoint events $A1$, $A2$, $\cdots$,An

$- \underbrace{P(A_1 \overset{\phi}{\cap} A_2)}_{0}$ } $P(\phi) = 0$

Notation:

- $P(A \cap B) = P(A \text{ and } B) = P(A, B)$,
- $P(A \cup B) = P(A \text{ or } B)$.

1.3.3 Finding Probabilities

We examine the events of interest and apply the axioms (and rules) of probabilities.

The following rules are direct consequences of the axioms of probability:

Example 1.10
Using the axioms of probability, prove the following:

a. For any event A, $P(A^c) = 1 - P(A)$.
b. The probability of the empty set is zero, i.e., $P(\emptyset) = 0$.
c. For any event A, $P(A) \leq 1$.
d. $P(A - B) = P(A) - P(A \cap B)$.
e. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, (inclusion-exclusion principle for $n = 2$).
f. If $A \subset B$ then $P(A) \leq P(B)$.

Solution

Take a look at the solutions !

(a) $P(A^c)$ $\Rightarrow$ $A \cup A^c = S$

$A \cap A^c = \phi$ 0

A, A^c are disjoint

$P(A \cup A^c) = P(A) + P(A^c) = P(S) = 1$

Note:

Inclusion-exclusion principle:

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

Generally for n events $A_1, A_2, \cdots, A_n$, we have

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j)$$
$$+ \sum_{i<j<k} P(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n-1} P\left(\bigcap_{i=1}^{n} A_i\right)$$

1.3.4 Discrete Probability Models

If S *(sample space)* is a *countable* set, we consider a **discrete** probability model.
If $A \subset S$ is an event, then A is also countable, and by the third axiom of probability we can write assuming that the outcomes (elements) of the random experiment are s_1, s_2, s_3, …. etc.

$$P(A) = P\left(\bigcup_{s_j \in A} \{s_j\}\right) = \sum_{s_j \in A} P(s_j).$$

To find probability of an event, all we need to do is sum the probabilities of individual elements (outcomes) in that set.

See Example 1.12

Finite Sample Spaces with Equally Likely Outcomes:

$S = \{s_1, s_2, \cdots, s_N\}$, where $\cancel{P(s_i) = P(s_j) \text{ for}}$ all $i, j \in \{1, 2, \cdots, N\}$.

$P(S) = P(s_1) + P(s_2) + \cdots + P(s_N) = 1$

all the same $P(s_i) = 1/N$ every i

In such a model, if A is any event with cardinality $|A|= M$, we can write

$$P(A) = \sum_{s_i \text{ in } A} P(s_i) = \underbrace{\frac{1}{N} + \frac{1}{N} + \frac{1}{N} + \cdots + \frac{1}{N}}_{M \text{ terms}} = \frac{M}{N} \begin{array}{l} \to \# \text{outcomes in } A \\ \to \# \text{outcomes in } S \end{array}$$

1.3.5 Continuous Probability Models

- Consider a scenario where your sample space S is, for example, [0,1].
- This is an uncountable set; we cannot list the elements in the set.
- Let's provide some intuition by looking at a simple example.

Example :

Your friend tells you that she will stop by your house sometime after or equal to 1p.m. and before 2p.m., but she cannot give you any more information as her schedule is quite hectic. Your friend is very dependable, so you are sure that she will stop by your house, but other than that we have no information about the arrival time. Thus, we assume that the arrival time is completely random in the 1p.m. and 2p.m. interval.
(As we will see, in the language of probability theory, we say that the arrival time

is "uniformly" distributed on the [1,2) interval).

Let T be the arrival time. (decimal) $1.5 \Leftrightarrow 1:30$

- What is the sample space S ?

  $S = \{t \mid 1 \le t < 2\}$

- What is the probability of $P(1.5)$? Why?

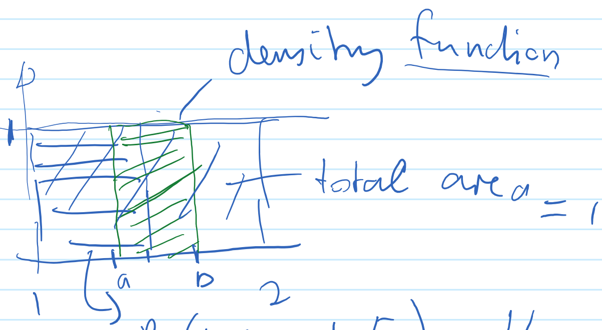
  " $P(T = 1.5)$ " $= 0$

  point

  interval

- What is the probability of $T \in [1, 1.5)$?

  $= .5$



$P_r(1 \le t < 1.5) = 1/2$

- For $1 \le a \le b \le 2$, what is $P(a \le T \le b) = P([a,b])$?

  $= (b - a)$ (area)

Our continuous probability model is a limit of a discrete probability model, when the precision becomes infinitely accurate. Thus, in reality we are always interested in the probability of some intervals rather than a specific point x.

In reality, the probability might not be uniform, so we need to develop tools that help us deal with general distributions of probabilities. These tools will be introduced in Chapter 4.

Discussion: *You might ask why $P(x)=0$ for all $x \in [1,2)$, but at the same time, the outcome of the experiment is always a number in [1,2) ?*