EE 525
Notes on Basic Probability
K. B. Kredo and F. W. DePiero

Probability: The basic entity treated by the theory of probability is called a “chance experiment.” It must always have an identifiable outcome belonging to a fixed, known set of possibilities.

Let ζ_i denote the outcome of a chance experiment. The whole set of possible outcomes is called the “universal set” and is denoted by S. The experiment itself may be designated by a letter E.

Examples:

1. The experiment E_a: A coin is tossed. The outcomes are $\zeta_1 = H$ and $\zeta_2 = T$.
   $S = (\zeta_1, \zeta_2) = (H, T)$
2. E_b: A coin is tossed three times in succession. Outcomes are triplets of heads or tails:
   $$\underset{\zeta_1}{HHH},\ \underset{\zeta_2}{HHT},\ \underset{\zeta_3}{HTH},\ \underset{\zeta_4}{HTT},\ \underset{\zeta_5}{THH},\ \underset{\zeta_6}{THT},\ \underset{\zeta_7}{TTH},\ \underset{\zeta_8}{TTT}$$
3. E_c: Two fair dice are thrown and the number of dots on the top of each is observed. There are 36 distinct outcomes that can be enumerated as $\underset{\zeta_1}{(1,1)}, \underset{\zeta_2}{(1,2)}, \underset{\zeta_3}{(1,3)}, \ldots, \underset{\zeta_{35}}{(6,5)}, \underset{\zeta_{36}}{(6,6)}$.
   Thirty-six, distinct 2-tuples describe the possible outcomes, and the sample space contains 36 elements.
4. E_d: The voltage at three different points a, b, and c of the circuit of a radio receiver is measured at the same time τ. The outcome is a triplet of numbers $\zeta = (V_a, V_b, V_c)$, which can be represented as a point in three-dimensional space, and S is the entire space.
5. E_e: The sequence of N digits (1’s and 0’s) emitted by a computer and sent over a communication line to a second computer is recorded during an interval of duration T. There are 2^N outcomes.

Events:

A collection of outcomes, ζ, of a chance experiment E forms a subset of the universal set S and is called an “event.” The event occurs whenever any outcome in it occurs.

Examples:

From Example 2,

$A =$ ”Heads comes first” $= \{\zeta_1, \zeta_2, \zeta_3, \zeta_4\}$

$B =$ ”The same face shows in the second and third tosses” $= \{\zeta_1, \zeta_4, \zeta_5, \zeta_8\}$

$C =$ ”Exactly two tails appear” $= \{\zeta_4, \zeta_6, \zeta_7\}$

From Example 5,

$A =$ ”The sequence of N digits contains j 0’s and N-j 1’s.”

$B =$ ”Each 0 is followed by k 1’s.”

The empty set, which contains no elements, is denoted by ϕ.

The experiment E_a has four events: $\phi, \{T\}, \{H\}, S$

In the experiment E_b each of the eight outcomes can be either included or excluded from the subset defining the event. There are 2^8 events. In a chance experiment with N outcomes 2^N different events can be defined. This is because there are two possibilities for each outcome (either exclude or include), and there are N outcomes total. Hence 2^N different events are possible.

Experiment E_d has an infinite number of events.

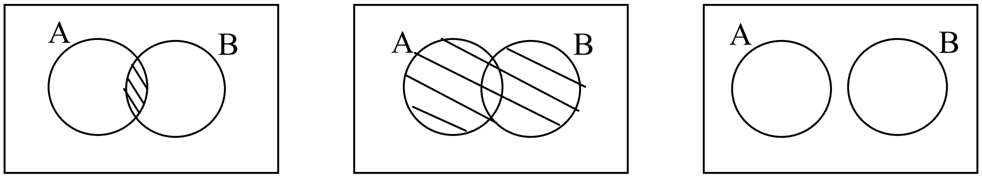
Set Definitions

A set is a collection of objects. The objects are called elements of the set. If a is an element of a set A, we can write $a \in A$. If a is not an element of A, we can write $a \notin A$.

A set is countable if its elements can be put in one-to-one correspondence with natural numbers. If a set is not countable it is called uncountable. A finite set is one that is either empty or has elements that can be counted. If a set is not finite it is called infinite. (Example: points on the real number line). An infinite set having countable elements is called countably infinite. (Example: the integers).

If every element of a set A is also an element in another set B, A is said to be contained in B, $A \subseteq B$. If at least one element exists in B, which is not in A, then A is a proper subset of B, $A \subset B$. Two sets A and B are called disjoint or mutually exclusive if they have no common elements.

Set Operations



$A \cap B$
(a)

$A \cup B$
(b)

(c)

Venn Diagrams

Relations among sets are often informatively represented by regions in a rectangle, which is called a Venn Diagram.

(a) Intersection: $D = A \cap B$. It is the set of all elements common to A and B.

$$D = \{\zeta : (\zeta \in A) \cap (\zeta \in B)\}$$

(b) Union: $C = A \cup B$. It is the set of all elements of A or B or both. $C = \{\zeta : (\zeta \in A) \cap (\zeta \in B)\}$

(c) For mutually exclusive sets A and B, $A \cap B = \phi$

The union and intersection of N sets A_n, n=1,2,3…N can be represented as:

$$C = A_1 \cup A_2 \cup \ldots \cup A_N = \bigcup_{n=1}^{N} A_n$$

$$D = A_1 \cap A_2 \cap \ldots \cap A_N = \bigcap_{n=1}^{N} A_n$$

The complement of a set A, denoted by $\overline{A}$, is the set of all elements not in A.

$$\overline{A} = \{\zeta : \zeta \notin A\} = S - A$$

Thus, $\overline{\phi} = S$, $\overline{S} = \phi$, $A \cup \overline{A} = S$, and $A \cap \overline{A} = \phi$.

Algebra of Sets

Commutative Law states that:

$$A \cap B = B \cap A$$

$$A \cup B = B \cup A$$

Distributive Law:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

Associative Law:

$$(A \cup B) \cup C = A \cup (B \cup C) = A \cup B \cup C$$

$$(A \cap B) \cap C = A \cap (B \cap C) = A \cap B \cap C$$

Probability and Axioms

Probabilities are numbers assigned to events in a way that is consistent with certain rules. The events are sets of outcomes of a chance experiment *E*. The rules are expressed in four axioms that form the basis of the mathematical theory of probability. We adopt the notation *P*(*A*) for "the probability of event *A*."

Let *A* be any event defined on a sample space.

Axiom I: $P(A) \geq 0$

Axiom II: $P(S) = 1$

Axiom III: If $A \cap B = \phi$, then $P(A \cup B) = P(A) + P(B)$ [Special case of Axiom IV]

Axiom IV: If $A_i \cap A_j = \phi \ \forall i, j, i \neq j$ then $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$

Corollary 1: $P(\overline{A}) = 1 - P(A)$

Proof:

$A \cap \overline{A} = \phi$ and $A \cup \overline{A} = S$

e.g., $P(A) + P(\overline{A}) = P(S) = 1$

Corollary 2: $0 \leq P(A) \leq 1$

Proof:

$P(A) \geq 0$ and $P(\overline{A}) \geq 0$ by Axiom I.
$P(\overline{A}) = 1 - P(A) \geq 0$ so $P(A) \leq 1$
e.g., $0 \leq P(A) \leq 1$

Corollary 3: If $A_i \cap A_j = \phi \;\; \forall i, j, i \neq j$ then $P\left(\bigcup_{i=1}^{N} A_i\right) = \sum_{i=1}^{N} P(A_i)$ (Note finite limit here)

Corollary 4: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Proof: Use a Venn Diagram.

Joint and Conditional Probability

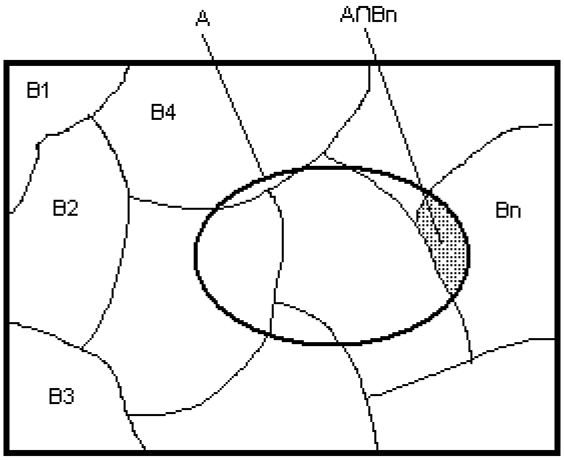
From the Venn diagrams above, it is clear that $P(A \cap B) = P(A) + P(B) - P(A \cup B)$ or $P(A \cup B) = P(A) + P(B) - P(A \cap B) \leq P(A) + P(B)$

Given some event *B* with nonzero probability, $P(B) > 0$, we define the conditional probability of an event *A*, given *B*, by

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

If *A* and *B* are mutually exclusive, then $A \cap B = \phi$ and $P(A|B) = 0$.

Total Probability



N mutually exclusive events B_n and another event *A*.

The probability of any event *A*, $P(A)$, defined on a sample space *S* can be expressed in terms of conditional probabilities.

Given *N* mutually exclusive events B_n, *n*=1,2,3…N ($B_n \cap B_m = \phi \;\; m \neq n = 1,2,3,\ldots,N$ and $\bigcup_{n=1}^{N} B_n = S$) the total probability of event *A* is given by

$$P(A)=\sum_{n=1}^{N}P(A|B_n)P(B_n)$$

Proof:

$$A\cap S = A = A\cap\left(\bigcup_{n=1}^{N}B_n\right)=\bigcup_{n=1}^{N}A\cap B_n$$

The events $A\cap B_n$ are mutually exclusive, so

$$P(A)=P(A\cap S)=P\left(\bigcup_{n=1}^{N}A\cap B_n\right)=\sum_{n=1}^{N}P(A\cap B_n)$$

Using conditional probability, $P(A)=\sum_{n=1}^{N}P(A|B_n)P(B_n)$

Independence

Two events are said to be statistically independent if:

$$P(A|B)=P(A),\ P(B)\neq 0 \text{ or } P(B|A)=P(B),\ P(A)\neq 0$$

or equivalently,

$$P(A\cap B)=P(A)P(B)$$

We already know that if the events are mutually exclusive then $P(A\cap B)=0$. If the events have nonzero probability of occurrence, then it is easily seen that two events can not be both mutually exclusive and statistically independent. For two events to be independent $A\cap B\neq\phi$.

Example:

In an experiment one card is selected from an ordinary 52-card deck. Define the events:

A – select a King
B – select a Jack or a Queen
C – select a Heart

The probabilities are: $P(A)=\frac{4}{52}$, $P(B)=\frac{8}{52}$, and $P(C)=\frac{13}{52}$.

$A\cap B=\phi$ so $P(A\cap B)=0$

$P(A\cap C)=\frac{1}{52}$ and $P(B\cap C)=\frac{2}{52}$

$P(A\cap B)=0\neq P(A)P(B)=\frac{32}{52^2}$ so A and B are not independent.

$P(A\cap C)=\frac{1}{52}=P(A)P(C)=\frac{1}{52}$ and $P(B\cap C)=\frac{2}{52}=P(B)P(C)=\frac{2}{52}$

so the pairs A,C and B,C are independent.

Bayes Theorem

$$P(B_n|A) = \frac{P(B_n \cap A)}{P(A)},\ P(A) \neq 0 \text{ and } P(A|B_n) = \frac{P(A \cap B_n)}{P(B_n)},\ P(B_n) \neq 0$$

$$P(A \cap B_n) = P(A|B_n)P(B_n) = P(B_n|A)P(A)$$

Bayes Theorem is derived from the above equations:

$$P(B_n|A) = \frac{P(A|B_n)P(B_n)}{P(A)}$$

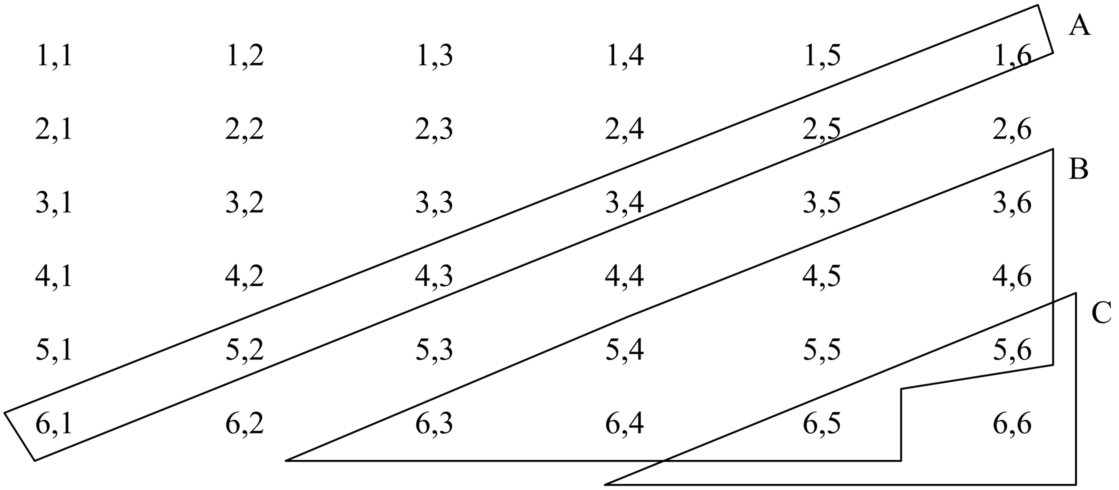
Another form of Bayes Theorem is obtained by substitution of

$P(A) = \sum_{n=1}^{N} P(A|B_n)P(B_n) = P(A|B_1)P(B_1) + ... + P(A|B_N)P(B_N)$ into the above equation.

$$P(B_n|A) = \frac{P(A|B_n)P(B_n)}{P(A|B_1)P(B_1) + ... + P(A|B_N)P(B_N)} = \frac{P(A|B_n)P(B_n)}{\sum_{i=1}^{N} P(A|B_i)P(B_i)}$$

Example

The experiment consists of observing the sum of numbers showing up when two dice are thrown. The set of all possible outcomes, or the sample space, consists of 36 points shown below. Each possible outcome corresponds to a sum having values from 2 to 12.



1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,2	6,3	6,4	6,5	6,6

Suppose we are interested in three events defined by the sum.

$A = \{\zeta : \zeta = 7\}$, $B = \{\zeta : 8 < \zeta \leq 11\}$, and $C = \{\zeta : 10 < \zeta\}$

There are 36 elementary events $A_{i,j}$ = { sum of outcomes (*i*,*j*) = *i* + *j* }

If the dice are fair: $P(A_{i,j}) = \frac{1}{36}$

$$P(A) = P\left(\bigcup_{i=1}^{6} A_{i,7-i}\right) = \sum_{i=1}^{6} P(A_{i,7-i}) = 6 \cdot \frac{1}{36} = \frac{1}{6}$$

$$P(B) = 9 \cdot \frac{1}{36} = \frac{1}{4} \qquad P(C) = 3 \cdot \frac{1}{36} = \frac{1}{12}$$

$$P(B \cap C) = 2 \cdot \frac{1}{36} = \frac{1}{18} \qquad P(B \cup C) = 10 \cdot \frac{1}{36} = \frac{5}{18}$$

Example:

Experiment E: Obtain a number by spinning the pointer on a fair wheel of chance that is labeled from 0 to 100. $S = \{x : 0 < x \le 100\}$ The probability of the pointer falling between any two numbers $x_1 \le x_2$ is $\frac{x_2 - x_1}{100}$.

To check this we see that $A = \{x : x_1 < x \le x_2\}$ satisfies Axiom I for all x_1 and x_2 and Axiom II when $x_2 = 100$ and $x_1 = 0$.

Suppose the periphery of the wheel is broken into N contiguous segments and $A_n = \{x : x_{n-1} < x \le x_n\}$ where $x_n = \frac{100n}{N}$, $n = 1, 2, 3, \dots, N$ and $x_0 = 0$.

If $P(A_n) = \frac{1}{N}$, then for any N $P\left(\bigcup_{n=1}^{N} A_n\right) = \sum_{n=1}^{N} P(A_n) = \sum_{n=1}^{N} \frac{1}{N} = 1 = P(S)$

Example:

In a box there are 100 resistors having resistance and tolerance as shown:

Resistance	**Tolerance**		**Total**
Ω	5 %	10 %	
22	10	14	24
47	28	16	44
100	24	8	32
Total	62	38	100

A resistor is chosen from the box with equal likelihood. Events are defined as:

A – 47Ω resistor B – 5% tolerance resistor C – 100Ω resistor

$$P(A) = \frac{44}{100} \qquad P(B) = \frac{62}{100} \qquad P(C) = \frac{32}{100}$$

$$P(A \cap B) = P(47\Omega and 5\%) = \frac{28}{100} \qquad P(A \cap C) = P(47\Omega and 100\Omega) = \frac{0}{100}$$

$$P(B \cap C) = P(5\% and 100\Omega) = \frac{24}{100}$$

$$P(A|B)=\frac{P(A\cap B)}{P(B)}=\frac{28/100}{62/100}=\frac{28}{62} \qquad P(A|C)=\frac{P(A\cap C)}{P(C)}=\frac{0/100}{32/100}=0$$

$$P(B|C)=\frac{P(B\cap C)}{P(C)}=\frac{24/100}{32/100}=\frac{24}{32}$$

Number of Combinations

The number of ways "N objects can be selected, k at a time" is denoted $\binom{N}{k}$. This is the number of combinations and is given by: $\binom{N}{k}=\frac{N!}{k!(N-k)!}$