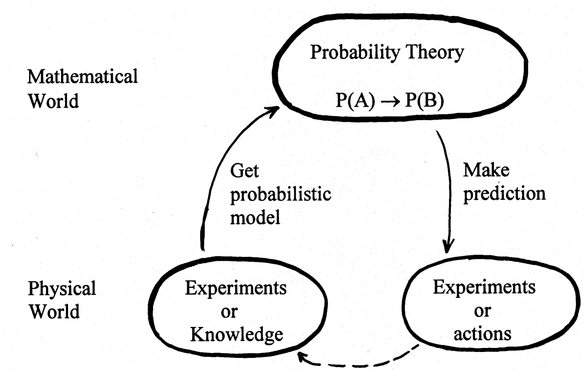
THEORY AND THE "REAL WORLD"



Examples

- Betting games
- Designing communication systems
- Detecting/filtering signals in noise
- Pattern recognition

The predictive power of the theory is based on the quality of the model.

Why do we need these models?

To handle ignorance or deterministic phenomena that are too difficult to model explicitly.

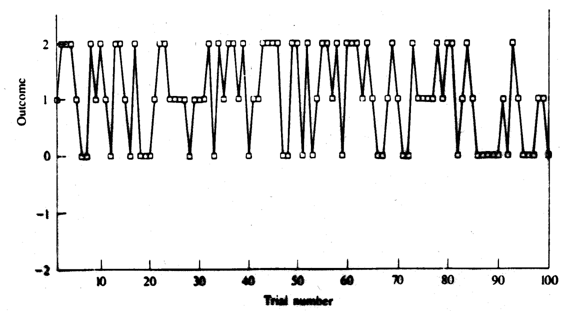
Random Experiment

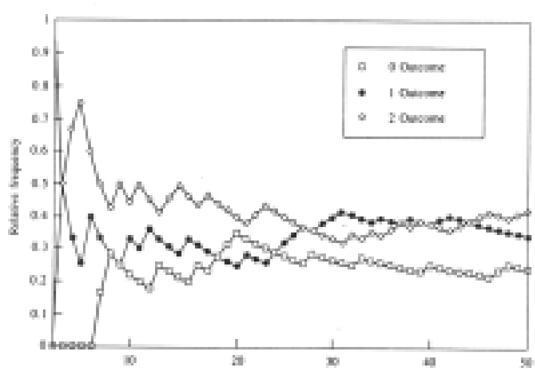
An experiment is said to be a random experiment if the outcome (of the experiment) varies in an unpredictable fashion when the experiment is repeated under the same conditions.

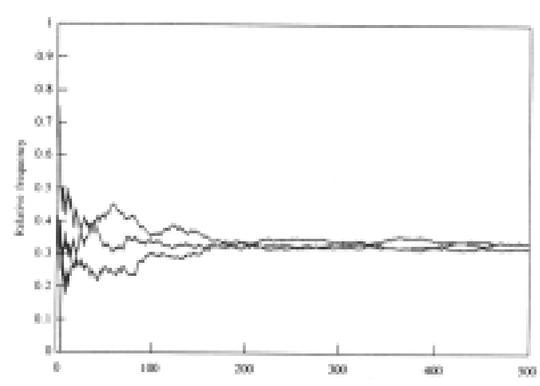
Example:

We have an urn containing three identical balls, labeled "0", "1", and "2". We first shake the urn to randomize the position of the balls and then select a ball from the urn. The number of the ball is noted and the ball is then returned to the urn.

The outcome of this experiment is a number from the set S = {0, 1, 2}. S is called the sample space.







An Experiment consists of:

1. The ***sample space*** S (the set of all possible outcomes).

   Consider the experiment of rolling a six-sided die. Each time we roll the die (a *trial*), we observe an *outcome*, ξ, e.g., $\xi = 5$.

   $$S = \{1,2,3,4,5,6\}$$

2. The set of all events.

   Subsets of S are called *events*. An event *occurs* if the outcome is contained within the event.

   Examples:

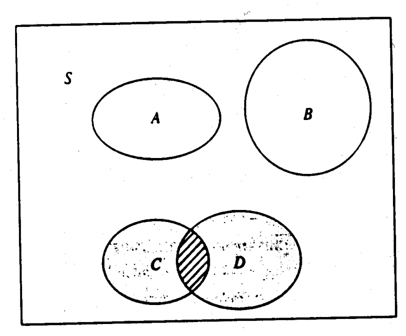
   S = the certain event
   ϕ = the impossible event
   {1}
   {2,4,6} = the "even" event
   {1,3,5} = the "odd" event

   Definitions: Let A and B be events. Then,

   - $A + B = A \cup B$ = A or B occurs.
   - $AB = A \cap B$ = both A and B occur.
   - A and B are *mutually exclusive* if whenever A occurs, B does not, i.e., $A \cap B = \phi$.

3. The probabilities of these events.

Mutually Exclusive (or Disjoint) Events



Which events are mutually exclusive?

Which are not?

AXIOMATIC VIEW OF PROBABILITY THEORY

To each event A, we assign a number P(A) called the ***probability*** of the event A which measures the likelihood of event A, such that the following axioms hold:

I. No event has negative probability: $P(A) \geq 0$

II. The certain event has probability 1: $P(S) = 1$

III. For any countable collection of events A1, A2, … which are mutually exclusive,

$$P(A_1 + A_2 + \cdots) = \sum_{k=1}^{\infty} P(A_k)$$

The rest of this course is based upon these three axioms.

Example 1

For the six-sided die, the probabilities of each outcome are usually chosen to be 1/6, i.e.,

P({1}) = P({2}) = P({3}) = P({4}) = P({5}) = P({6}) = 1/6.

Consequently,

P(“even”) = P({2,4,6}) = P({2}) + P({4}) + P({6}) = 1/2.

P(“less than three”) = P({1,2}) = 1/3

P(“even or less than three”) = P({1,2,4,6}) = 2/3

BASIC PROPERTIES

1. $P(\phi) = 0$.

   Proof: $A = A + \phi \Rightarrow P(A + \phi) = P(A)$.
   Since A and ϕ are mutually exclusive, by axiom 3,
   $$P(A) = P(A) + P(\phi) \Rightarrow P(\phi) = 0.$$

2. For any event A, $P(A) = 1 - P(\overline{A}) \leq 1$.

   Proof: Since $A + \overline{A} = S$,
   $$P(A + \overline{A}) = P(S)$$
   and by axioms 3 and 2
   $$P(A) + P(\overline{A}) = 1.$$
   By axiom 1,
   $$P(A) = 1 - P(\overline{A}) \leq 1.$$

3. For any events A and B,
   $$P(A + B) = P(A) + P(B) - P(AB).$$

4. For any events A and B such that $A \subset B$,
   $$P(A) \leq P(B).$$

Example 2

P(“even or less than three”)

= P(“even”)+P(“less than three”)-P(“even and less than three”)

$= \frac{1}{2} + \frac{1}{3} - \frac{1}{6} = \frac{2}{3}$.

RANDOM EVENTS

- An event is a subset of the sample space and consists of one or more outcomes (simple or compound).

  Example: When two dice are rolled, the event E={“sum is 10”} consists of the 3 outcomes (5,5), (6,4), (4,6).

  The probability of any event is equal to the sum of the probabilities of the outcomes that make up the event.

- When we conduct a random experiment, we can use set notation to describe possible *outcomes*.

- **Complementary event:** If A is an event, the set of all outcomes that are not in A is called the complement of A

  Example: Roll a six-sided die.

  Possible Outcomes: S = {1,2,3,4,5,6}

  An event is any subset of possible outcomes: A = {1,2}

  The *complementary event*: $\overline{A} = S - A = \{3,4,5,6\} = A^c$

  The set of all outcomes is the *certain event*: S

  The *null event*: ϕ

  Example: If S = {a,b,c}, there are 8 events:
  ø (empty set), {a}, {b}, {c}, {a,b}, {a,c}, {b,c}, {a,b,c} = S

  The sample space S is called the *sure event* or the *certain event* – it always occurs

  Events with one outcome are called *elementary events*

Example:

An urn contains 6 red balls and 4 white balls. A ball is drawn at random. What is the probability that it is red?

Comment: “random” generally means that the classical approach is to be used and all outcomes are equally likely.

There are ten different outcomes of this experiment (not two!) and the sample space is

S= {R1, R2, R3, R4, R5, R6, W1, W2, W3, W4}

and not S= {R, W}.

P(Red Ball) = 6/10

- What if the experiment consists of drawing two balls at random?

  Important question to be decided: How are these balls drawn? Is the first one drawn put back in the urn before the second is picked or not? Or are we reaching in and grabbing two at once?

➢ **Sampling with replacement:** first ball is replaced in urn before the second ball is drawn.

S has 100 elements that are *vectors* of the form (X,Y) with X and Y taking on all possible values from the set

{R1, R2, R3, R4, R5, R6, W1, W2, W3, W4}

If (X,Y) = (R1,W3), this means the first ball drawn was R1 and the second ball drawn wasW3.

It is possible for (X,Y) to be (say) the vector (R3,R3) indicating that the same red ball R3 was drawn the second time as well.

1. P(both balls drawn are red) = 6x6/100 = 9/25
2. P(both balls drawn are white) = 4x4/100 = 4/25
3. P(one red and one white) = 1 – 9/25 – 4/25 = 12/25

   More directly, we get the answer as 6x4/100 + 4x6/100 = 48/100 = 12/25
4. P(first ball is red) = 6/10

➢ **Sampling without replacement**: If the first ball drawn is not replaced before the second is drawn, the sample space has 10x9 = 90 elements that are *vectors* of the form (X,Y) with X and Y taking on all possible *different* values from the set

{R1, R2, R3, R4, R5, R6, W1, W2, W3, W4}

If (X,Y) = (R1,W3), this means the first ball drawn was R1 and the second ball drawn was W3.

Note that it is *not* possible for (X,Y) to be the vector (R3,R3) because R3 is not being put back into the urn before the second ball is drawn.

1. P(both balls drawn are red) = 6x5/90 = 1/3
2. P(both balls drawn are white) = 4x3/90 = 2/15
3. P(one red and one white) = 1 – 1/3 – 2/15 = 8/15

   More directly, we get the answer as 6x4/90 + 4x6/90 = 48/90 = 8/15
4. P(first ball is red) = 6/10

➢ If we reach in and grab two at the same time, we get a (randomly chosen) subset of size 2 from a set of size 10

Now, each element of S is *a set* of the form {X,Y} with X and Y taking on all possible *distinct* values from the set

{R1, R2, R3, R4, R5, R6, W1, W2, W3, W4}

<u>Note:</u> Previously, the elements of S were vectors (X,Y) where X was the first ball drawn and Y the second, and (R3,W1) was a different outcome from (W1,R3). Now, the elements of S are *sets* {X,Y} and we no longer can talk about the first ball drawn and the second ball drawn; both balls are drawn at the same time, and {R3,W1} is the same set as {W1,R3}.

S has $\binom{10}{2} = 45$ elements

1. P(both balls drawn are red) = $\dfrac{\binom{6}{2}}{\binom{10}{2}} = \dfrac{6x5}{10x9} = \dfrac{1}{3}$
2. P(both balls drawn are white) = $\dfrac{\binom{4}{2}}{\binom{10}{2}} = \dfrac{4x3}{10x9} = \dfrac{2}{15}$

P(one red and one white) = 1 – 1/3 – 2/15 = 8/15

Example: Roll a fair die once.

S =

The die turns up even

A =

The die turns up less than 5

B =

The die turns up greater than 4

C =

The die turns up 6

D =

Let's build a probability measure:

$P(A) =$

$P(B) =$

$P(C) =$

$P(D) =$

$P(\phi) =$

$P(B \cup C) = P(B) + P(C)$

$P(B \cup D) = P(B) + P(D)$

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

Example:

A survey of staff leaving a company indicated the following reasons for leaving:

- dissatisfaction with salary: 30%
- dissatisfied with tasks they had to do: 20%
- 12% said both salary and dissatisfaction with tasks were the reason for leaving

What is the probability that a person leaving work is dissatisfied with salary or with the tasks, or with both?

Let S = {"salary dissatisfaction"}
T = {"dissatisfaction with tasks"}

Given: P(S) = 0.30
P(T) = 0.20
$P(S \cap T) = 0.12$

$$P(S \cup T) = P(S) + P(T) - P(S \cap T)$$
$$= 0.3 + 0.2 - 0.12$$
$$= 0.38 (or\ 38\%)$$

EXAMPLE: An experiment consists of observing the sum of the numbers showing up when 2 dice are thrown.

⇒ The sample space S consists of $6^2 = 36$ points.

⇒ Each possible outcome corresponds to a sum having values from 2 to 12.

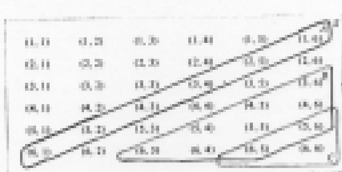
⇒ Consider these events:

⇒ $A = \{\text{"sum = 7"}\}$

⇒ $B = \{\text{"8 < sum ≤ 11"}\}$

⇒ $C = \{\text{"sum > 10"}\}$

$A_{ij} = \{\text{"sum for outcome } (i,j) = i+j\text{"}\}$ — ELEMENTARY EVENT



Since the dice are assumed fair, $P(A_{ij}) = \frac{1}{36}$.

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

$P(B) = 9\left(\frac{1}{36}\right) = \frac{1}{4}$

$P(C) = 3\left(\frac{1}{36}\right) = \frac{1}{12}$

$P(BC) = 2\left(\frac{1}{36}\right) = \frac{1}{18}$

$P(B+C) = 10\left(\frac{1}{36}\right) = \frac{5}{18}$

CONDITIONAL PROBABILITY

In many cases, we may have (through observation) partial knowledge of the outcome of an event.

<u>Definition</u>: The ***conditional probability*** of an event A with respect to an event M is defined as the ratio

$$P(A|M) = \frac{P(AM)}{P(M)}$$

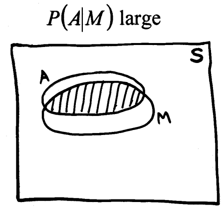
where it is assumed that $P(M) \neq 0$.

<u>Note</u>: P(A) is often called the ***<u>a priori</u>*** probability. A priori means relating to reasoning from self-evident propositions or presupposed by experience.

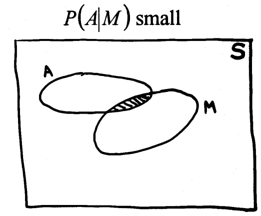
P(A|M) is often called the ***<u>a posteriori</u>*** probability. A posteriori means relating to reasoning from observed facts.

<u>Interpretation</u>

$P(A|M)$ large



$P(A|M)$ small



MULTIPLICATION RULE

- probability that events A and B will both occur
- a rearrangement of the conditional probability equation
- intersection also called joint probability

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

Example:

1. A box contains 4 red balls, 6 green balls, and 3 yellow balls. A ball is selected at random. Find the probability that the selected ball is green.

P(Green) = 6 / 13

2. A second ball is selected without replacement. Find the probability that:

- It is green (if the first ball was also green)

P(G|G) = 5 / 12

- The first is green and the second is red

$$P(G\cap R)=P(G)\cdot P(R|G)=\frac{6}{13}\cdot\frac{4}{12}=\frac{2}{13}$$

STATISTICAL INDEPENDENCE

A and B are said to be *statistically independent* when the occurrence of one is not affected by the occurrence of the other.

In this case, the conditional probability is equal to the unconditional probability, i.e.,

$$P(A|B)=P(A)$$

Consequently,

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

PROPERTIES OF CONDITIONAL PROBABILITY

1. If $M\subset A$, then $P(A|M)=1$.
2. If $A\subset M$, then $P(A|M)=\frac{P(A)}{P(M)}\geq P(A)$.
3. For fixed M, $P(A|M)$ is a probability measure.

Proof:

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

Follows since $P(AM)\geq 0$ and $P(M)>0$.

Axiom II: $P(S|M)=1$

Since SM = M, $P(S|M)=\frac{P(SM)}{P(M)}=\frac{P(M)}{P(M)}=1$.

Axiom III: If A and B are mutually exclusive, then

$$P(A+B|M)=P(A|M)+P(B|M).$$

By definition,

$$P(A+B|M)=\frac{P((A+B)M)}{P(M)}=\frac{P(AM+BM)}{P(M)}.$$

Since A and B are mutually exclusive, so are AM and BM. Hence,

$$P(A+B|M)=\frac{P(AM)}{P(M)}+\frac{P(BM)}{P(M)}=P(A|M)+P(B|M).$$

Example 3

Consider a box with 3 white balls and 2 red balls. Two balls are picked in succession. What is the probability that the first ball is white and the second ball is red?

Define the events

W1 = {“white picked first”}

R2 = {“red picked second”)

Since $P(R_2|W_1)=\frac{P(R_2W_1)}{P(W_1)}$, $P(W_1R_2)=P(R_2|W_1)P(W_1)$.

Since $P(R_2|W_1)=\frac{2}{4}$ and $P(W_1)=\frac{3}{5}$, $P(W_1R_2)=\frac{2}{4}\cdot\frac{3}{5}=\frac{6}{20}=\frac{3}{10}$.

Example 4

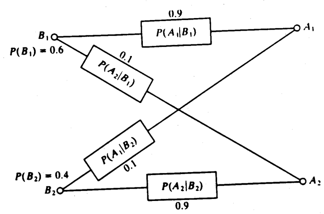
What is the probability that a white ball was picked first, given that a red ball was picked second?

By definition, $P(W_1|R_2)=\frac{P(W_1R_2)}{P(R_2)}$. To find $P(R_2)$, we enumerate all possibilities:

$$\left.\begin{array}{l}\#\text{ of outcomes in } R_1R_2=2\cdot 1=2\\ \#\text{ of outcomes in } W_1R_2=3\cdot 2=6\end{array}\right\}8\text{ total}$$

Thus, $P(W_1|R_2)=\frac{6/20}{8/20}=\frac{3}{4}>\frac{3}{5}$ (the *a priori* probability)

EXAMPLE: Binary Symmetric Communication System



$B_1 = \{$"the symbol transmitted is 1"$\}$
$B_2 = \{$"the symbol transmitted is 0"$\}$
$A_1 = \{$"the symbol received is 1"$\}$ ⟩ mutually exclusive events
$A_2 = \{$"the symbol received is 0"$\}$

➡ "Received" symbol probabilities:

$$P(A_1) = P(A_1|B_1)P(B_1) + P(A_1|B_2)P(B_2) = 0.9(0.6) + 0.1(0.4) = 0.58$$

$$P(A_2) = P(A_2|B_1)P(B_1) + P(A_2|B_2)P(B_2) = 0.1(0.6) + 0.9(0.4) = 0.42$$

$0.58 + 0.42 = 1$

➡ $P(B_1|A_1) = \dfrac{P(A_1|B_1)P(B_1)}{P(A_1)} = \dfrac{0.9(0.6)}{0.58} = \dfrac{0.54}{0.58} \simeq 0.931$

$P(B_2|A_2) = \dfrac{P(A_2|B_2)P(B_2)}{P(A_2)} = \dfrac{0.9(0.4)}{0.42} = \dfrac{0.36}{0.42} \simeq 0.857$

(probabilities of correct transmission)

$P(B_1|A_2) = \dfrac{P(A_2|B_1)P(B_1)}{P(A_2)} = \dfrac{0.1(0.6)}{0.42} = \dfrac{0.06}{0.42} \simeq 0.143$

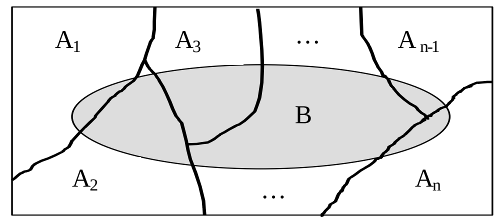
$P(B_2|A_1) = \dfrac{P(A_1|B_2)P(B_2)}{P(A_1)} = \dfrac{0.1(0.4)}{0.58} = \dfrac{0.04}{0.58} \simeq 0.069$

(probabilities of system error)

TOTAL PROBABILITY THEOREM

Definition: A finite or countable collection of sets $A_1, A_2, A_3, \ldots, A_n$ is called a ***partition*** of a set S if

1. They are mutually exclusive.
2. $\bigcup_{i=1}^{n} A_i = S$.



Theorem: If $A_1, A_2, A_3, \ldots, A_n$ are events which form a partition of the sample space S and B is an arbitrary event, then

$$P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i).$$

Proof: $B = BS = B(A_1 + A_2 + \cdots + A_n) = BA_1 + BA_2 + \cdots + BA_n$.
Since the $BA_1, BA_2, \ldots, BA_n$ are mutually exclusive, then by axiom 3

$$P(B) = \sum_{i=1}^{n} P(BA_i).$$

The theorem follows since $P(BA_i) = P(B|A_i)P(A_i), \ \forall i$.

BAYES' THEOREM (OR RULE)

Let the events $A_1, A_2, A_3, \ldots, A_n$ form a partition of S and B be an arbitrary event such that $P(B) > 0$. Suppose that event B occurs. The probability of event A_i occurring is given by

$$P(A_i|B) = \frac{P(BA_i)}{P(B)} = \frac{P(B|A_i)P(A_i)}{\sum_{k=1}^{n} P(B|A_k)P(A_k)}.$$

This is called Bayes' rule.

Example 5

Consider a box with 3 white balls and 2 red balls. Two balls are picked in succession. What is the probability that a white ball was picked first given that a red ball was picked second?

Define: W1 = {"white picked first"}
R2 = {"red picked second"}
R1 = {"red picked first"}

By Bayes' rule,

$$P(W_1|R_2) = \frac{P(R_2|W_1)P(W_1)}{P(R_2|W_1)P(W_1) + P(R_2|R_1)P(R_1)} = \frac{\left(\frac{1}{2}\right)\left(\frac{3}{5}\right)}{\left(\frac{1}{2}\right)\left(\frac{3}{5}\right) + \left(\frac{1}{4}\right)\left(\frac{2}{5}\right)} = \frac{3}{4}$$

Example 6

We have 3 "identical" cards whose sides are colored as follows:

RED/RED BLACK/BLACK RED/BLACK

A card is chosen "at random" and placed on the table. If the upper side of the card is red, what is the probability that the other side is black?

Define: RR = {"RED/RED card is chosen"}
RB = {"RED/BLACK card is chosen"}
BB = {"BLACK/BLACK card is chosen"}
R = {"Upper side of chosen card is RED"}

By Bayes' theorem,

$$P(RB|R) = \frac{P(R|RB)P(RB)}{P(R|RR)P(RR) + P(R|RB)P(RB) + P(R|BB)P(BB)} = \frac{\left(\frac{1}{2}\right)\left(\frac{1}{3}\right)}{(1)\left(\frac{1}{3}\right) + \left(\frac{1}{2}\right)\left(\frac{1}{3}\right) + (0)\left(\frac{1}{3}\right)} = \frac{1}{3}$$

INDEPENDENT EVENTS

Definition: Two events A and B are said to be ***independent*** if $P(AB) = P(A)P(B)$.

Knowledge of event B occurring "says nothing" about whether or not event A occurred:

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

Interpretation:

$$\frac{P(AB)}{P(B)} = \frac{P(A)}{P(S)}$$

Lemma: If A and B are independent events, then A and $\overline{B}$ are also independent events.

Proof: $A = SA = (B + \overline{B})A = AB + A\overline{B}$

$$P(A) = P(AB) + P(A\overline{B})$$

$$P(A\overline{B}) = P(A) - P(AB) = P(A) - P(A)P(B)$$
$$= P(A)(1 - P(B)) = P(A)P(\overline{B})$$

Example 7

Suppose we toss two fair dice. Let A = {"sum of dice equals 6"} and B = {"first die equals 4"}. Are A and B independent events?

$$P(AB) = P(\{(4,2)\}) = \frac{1}{36}$$

$$P(A)P(B) = \left(\frac{5}{36}\right)\left(\frac{1}{6}\right) = \frac{5}{216}$$

Since $P(AB) \neq P(A)P(B)$, A and B are **not** independent events.

Example 8

Suppose we toss two fair dice. Let A = {"sum of dice equals 7"} and B = {"first die equals 4"}. Are A and B independent events?

$$P(AB) = P(\{(4,3)\}) = \frac{1}{36}$$

$$P(A)P(B) = \left(\frac{6}{36}\right)\left(\frac{1}{6}\right) = \frac{1}{36}$$

Since $P(AB) = P(A)P(B)$, A and B are independent events.

INDEPENDENCE OF 3 EVENTS

Definition: Events A1, A2, and A3 are said to be ***independent*** if

1. They are pair-wise independent, i.e.,
$$P(A_i A_j) = P(A_i)P(A_j), \; i \neq j$$
2. $P(A_1 A_2 A_3) = P(A_1)P(A_2)P(A_3)$

Note: It is not sufficient to check only pair-wise independence. Say we toss two fair dice. Consider the following three events:

A = {"first die equals 3"}
B = {"sum of dice equals 7"}
C = {"second die equals 2"}

Although $P(A) = P(B) = P(C) = \frac{1}{6}$

and $P(AB) = P(BC) = P(AC) = \frac{1}{36}$,

the three events are not independent since

$$P(ABC) = 0 \neq P(A)P(B)P(C).$$

Independence of n Events

Assume that for all k < n we have defined independence for k events. The events $A_1, A_2, \ldots, A_n$ are said to be independent if

1. Any group of k of them are independent for any k < n
2. $P(A_1 A_2 \cdots A_n) = \prod_{i=1}^{n} P(A_i)$.