**Charles Boncelet, "Probability, Statistics, and Random Signals,"**
**Oxford University Press, 2016. ISBN: 978-0-19-020051-0**

# Chapter 1: PROBABILITY BASICS

**Sections**



Notes and figures are based on or taken from materials in the course textbook: Charles Boncelet, *Probability, Statistics, and Random Signals*, Oxford University Press, February 2016.

*An understanding of Probability and Statistics is necessary in most if not all work related to science and engineering.*

*Statistics*: the study of and the dealing with data.

*Probability*: the study of the likeliness of result, action or event occurring.

Often based on prior knowledge or the statistics of similar or past events!

Terms: *Random Variables, Random Processes or Stochastic Processes*

For any **measured phenomenon** there will be Uncertainty, Expected Variations, Randomness, or even Expected Errors included.

- when an outcome is non-deterministic
- where an exact value is subject to errors … e.g. noise, measurement noise and error

Easy examples of such phenomenon include all games of chance

- Flipping coins, rolling dice, dealing cards, etc.

Engineering Applications include

- Realistic signals – with noise or characteristic “unknown” parts
- Signal-to-noise Ratios, Noise-Power Measurements, Background Noise
- Expected Values, Variances, Distributions
- Thermal Motion, Electron Movement
- Reliability, Quality, Failure Rates, etc.

Probability theory is necessary for engineering system modeling and simulations.

- unknown initial conditions (random)
- noisy measurements, expected inaccuracies, etc. during operation

## Different Kinds of Probability

Suggested that there are essentially 4 types

- Probability by Intuition
    - "Lucky Numbers"

- Probability as the Ratio of Favorable to Total Outcomes (Classical Theory)
    - Measured Statistical Expectations

- Probability as a Measure of the Frequency of Outcomes

- Probability Based on Axiomatic Theory
    - Theory and mathematical derivations.
    - Should be validated by experimentation.

## Definitions used in Probability

Experiment

- An experiment is some action that results in an outcome.
- A random experiment is one in which the outcome is uncertain before the experiment is performed.

Possible Outcomes

- A description of all possible experimental outcomes.
- The set of possible outcomes may be discrete or form a continuum.

Trials

- The single performance of a well-defined experiment.

Event

- An elementary event is one for which there is only one outcome.
- A composite event is one for which the desired result can be achieved in multiple ways. Multiple outcomes result in the event described.

Equally Likely Events/Outcomes

- When the set of events or each of the possible outcomes is equally likely to occur.
- A term that is used synonymously to equally likely outcomes is a uniform random variable.

## Probability as the Ratio of Favorable to Total Outcomes (Classical Theory) – 2 Dice example

**Table 1.2-1** Outcomes of Throwing Two Dice

| 2nd die | 1st die 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| 6 | 7 | 8 | 9 | 10 | 11 | 12 |



Probability, Statistics and Random Processes for Engineers, 4th ed., Henry Stark and John W. Woods, Pearson Education, Inc., 2012.

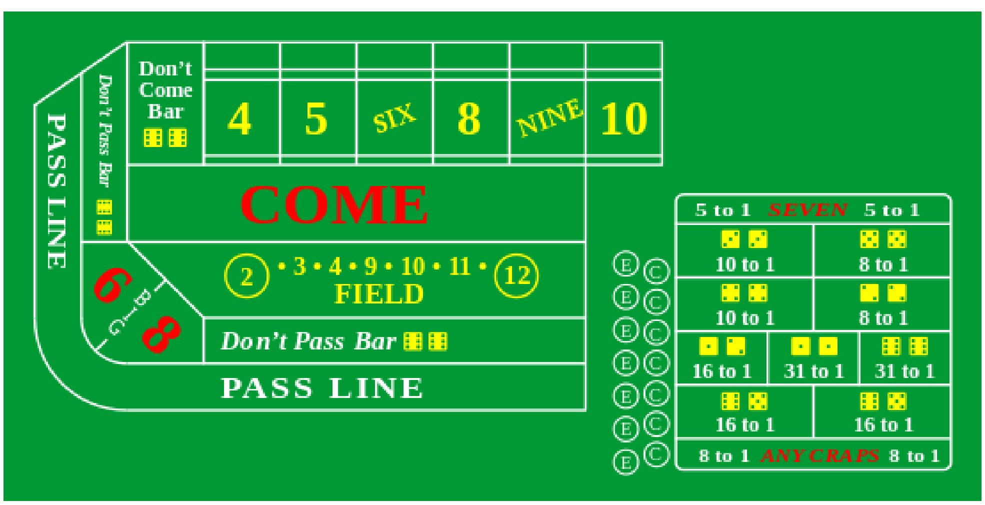
An elemental event can be defined as the total of the two die …

There are 36 possible outcomes

The total number of outcomes resulting in each unique event is known.

The probability of each event can be computed and described
… if the die are "fair".

So … the "true odds" can be computed … and a gambling game with skewed odds in "the houses" favor can be created …




from: https://en.wikipedia.org/wiki/Craps

Notes and figures are based on or taken from materials in the course textbook: Charles Boncelet, *Probability, Statistics, and Random Signals*, Oxford University Press, February 2016.

## Probability as the Ratio of Favorable to Total Outcomes (Classical Theory) – flipping two coins example

Flip two coins:

What are the possible outcomes {HH, HT, TH, TT}

Define an event as the getting of ***at least one Tail***.

Probability is the favorable outcomes/total outcomes = 3/4

Possible Outcomes with Probabilities:

HH – probability ¼ or 25%

HT – probability ¼ or 25%

TH – probability ¼ or 25%

TT – probability ¼ or 25%

Possible events: one head, one tail, at least one head, at least one tail, at most one head, at most one tail, two heads, two tails, no heads, or no tails.

The events can be a single possible outcome or a combination of multiple outcomes.

As a note … "not possible outcome(s)" can be defined …. a "null" event. (prob. 0%)

(e.g. three heads with two coins)

## Probability as a Measure of the Frequency of Outcomes

Experiment: Selecting a sequence of random numbers.

- The random numbers are between 1 and 100.

Determining the relative frequency of a single number as from 01 to 10,000 numbers are selected.

- The statistics of "observed events" is approaching 1/100 …. (infinite trials?)

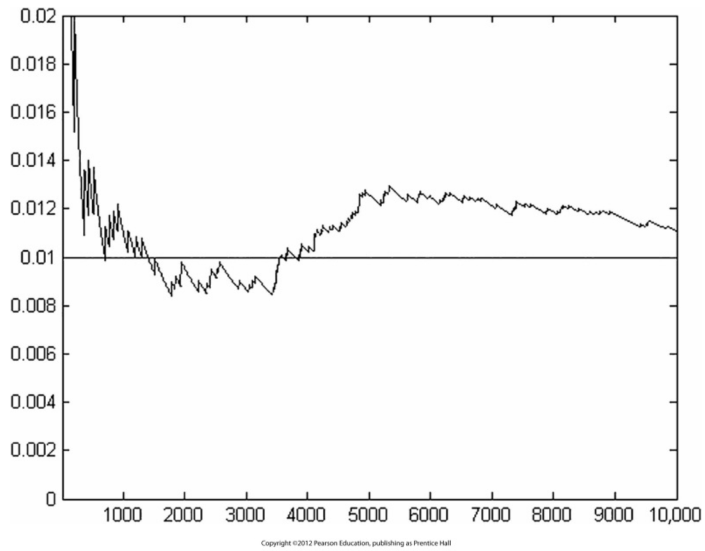



**Figure 1.2-1 Event = {occurrence of number 5}**

(Numbers derived from website RANDOM.ORG).

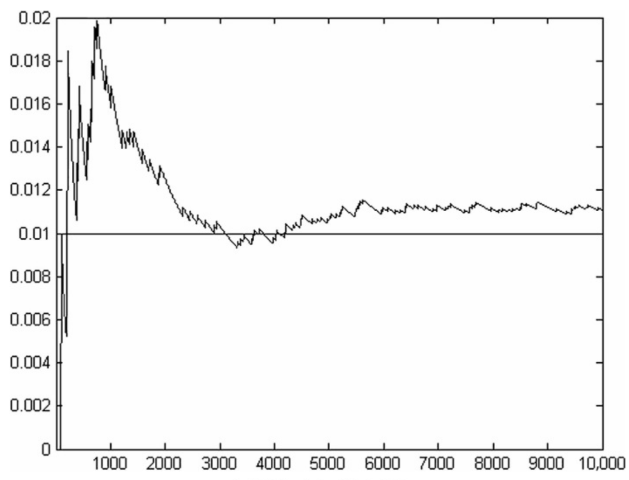



**Figure 1.2-2 Event = {occurrence of number 23}**

Probability, Statistics and Random Processes for Engineers, 4th ed., Henry Stark and John W. Woods, Pearson Education, Inc., 2012.

## Probability Based on an Axiomatic Theory

Develop the coherent mathematical theory:

- Statistics collected data on random experiments
  - Possible outcomes, sample space, events, etc.

- From the statistics, probability structure can be observed and defined
  - Random processes follow defined probabilistic models of performance.

- Mathematical properties applied to probability derives new/alternate expectations
  - Probabilistic expectations can be verified by statistical measurement.

*This can be considered as modeling a system prior to or instead of performing an experiment. Note that the results are only as good as the model or "theory" matching the actual experiment.*

## Misuses, Miscalculations, and Paradoxes in Probability

Old time quotation … "There are three kinds of lies: lies, damned lies, and statistics!"
https://en.wikipedia.org/wiki/Lies,_damned_lies,_and_statistics

From the CNN headlines …

"Math is racist: How data is driving inequality", by Aimee Rawlins, September 6, 2016
http://money.cnn.com/2016/09/06/technology/weapons-of-math-destruction/index.html

*Think of a "blind" allocation system vs. a statistically derived allocation system. Who are the winners and the losers ? Were some "losers" "statistically included or excluded"?*

Another example …

As a person, you are a unique individual and not a statistical probability …

but future "chances" may be based on others like you that have come before.

The class as a whole may exhibit statistical expectations … although it is made up of unique individuals.

For Sci-Fi readers ….

Issac Asimov's Foundation Trilogy – "psychohistory" used to predict the future ….

https://en.wikipedia.org/wiki/Psychohistory_(fictional)

Notes and figures are based on or taken from materials in the course textbook: Charles Boncelet, *Probability, Statistics, and Random Signals*, Oxford University Press, February 2016.

## Sets, Fields and Events

Conceptually Defining a Problem

- Relative Frequency Approach (statistics)
- Set Theory Approach (formal math)
- Venn Diagrams (pictures based on set theory)

***Find the best method that works for you to conceptualize probability.***

If you like "pictures" try to use Venn Diagrams to help understand the concepts.

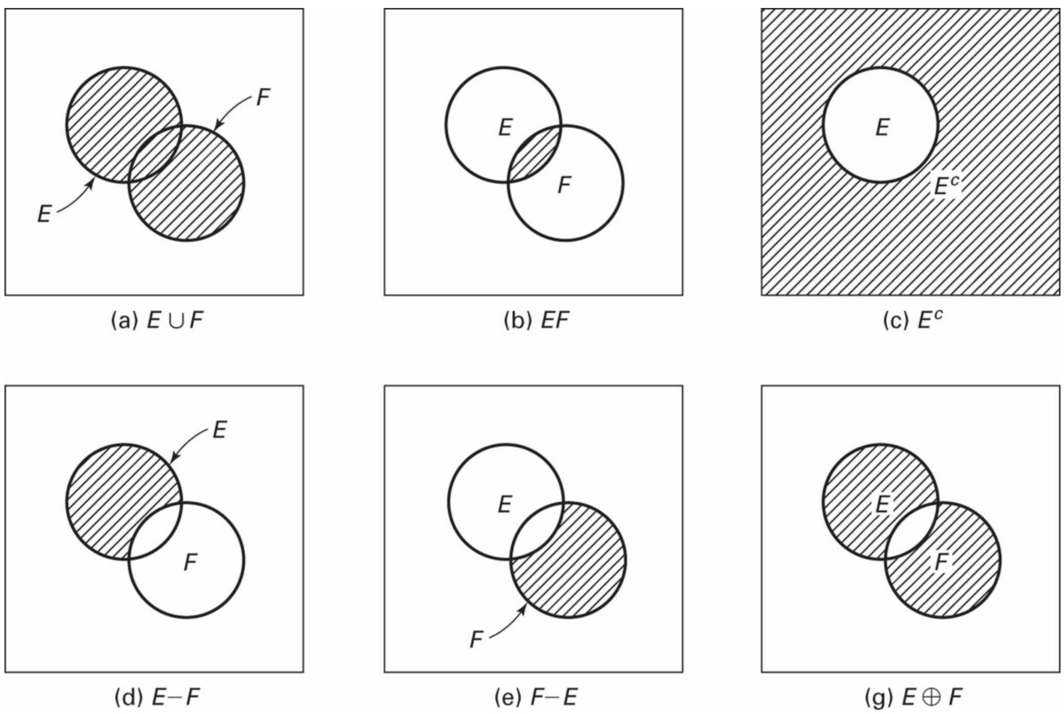




Figure 1.4-1 Venn diagrams for set operations.

Probability, Statistics and Random Processes for Engineers, 4th ed., Henry Stark and John W. Woods, Pearson Education, Inc., 2012.

# Set Theory Definitions – (A review?!)

Set

- A collection of objects known as elements

$$A = \{a_1, a_2, \cdots, a_n\}$$

Subset

- The set whose elements are all members of another set (usually larger but possible the same size).

$$B = \{a_1, a_2, \cdots, a_{n-k}\} \text{ therefore } B \subset A$$

Space

- The set containing the largest number of elements or all elements from all the subsets of interest. For probability, the set containing the event description of all possible experimental outcomes.

$$A_i \subset S \text{, for all i subsets}$$

Null Set or Empty Set

- The set containing no elements … $A \subset \varnothing$

Venn Diagrams can help when considering set theory …

- A graphical (geometric) representation of sets that can provide a way to visualize set theory and probability concepts and can lead to an understanding of the related mathematical concepts.

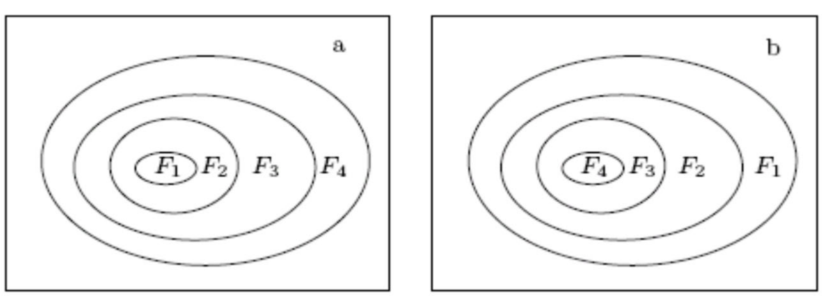



**Figure 2.2** (a) Increasing sets, (b) decreasing sets

from: Robert M. Gray and Lee D. Davisson, *An Introduction to Statistical Signal Processing*, Cambridge University Press, 2004. A pdf file version can be found at http://www-ee.stanford.edu/~gray/sp.html

Notes and figures are based on or taken from materials in the course textbook: Charles Boncelet, *Probability, Statistics, and Random Signals*, Oxford University Press, February 2016.

# More Set Theory Definitions

Equality

- Set A equals set B if and only if (iff) every element of A is an element of B AND every element of B is an element of A.

$$A = B \ iff \ \ A \subset B \ \ and \ \ B \subset A$$

Sum or Union (logic OR function)

- The sum or union of sets results in a set that contains all of the elements that are elements of every set being summed.

$$S = A_1 \cup A_2 \cup A_3 \cdots \cup A_N$$

- Laws for Unions

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

$$A \cup A = A$$

$$A \cup \varnothing = A$$

$$A \cup S = S$$

$$A \cup B = A, \ if \ \ B \subset A$$

Products or Intersection (logic AND function)

- The product or intersection of sets results in a set that contains all of the elements that are present in every one of the sets.

$$S \cap \varnothing = \varnothing$$

- Laws for Intersections

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

$$A \cap A = A$$

$$A \cap \varnothing = \varnothing$$

$$A \cap S = A$$

$$A \cap B = B, \ if \ \ B \subset A$$

Mutually Exclusive or Disjoint Sets

- Mutually exclusive or disjoint sets of no elements in common.

$$A \cap B = \varnothing$$

- NOTE: The intersection of two disjoint sets <u>is a set</u> … the null set!

Notes and figures are based on or taken from materials in the course textbook: Charles Boncelet, *Probability, Statistics, and Random Signals*, Oxford University Press, February 2016.

Complement

- The complement of a set is the set containing all elements in the space that are not elements of the set.

$$A \cap \overline{A} = \varnothing \text{ and } A \cup \overline{A} = S$$

- Laws for Complement

$$\overline{\varnothing} = S$$

$$\overline{S} = \varnothing$$

$$\overline{\left(\overline{A}\right)} = A$$

$$\overline{A} \subset \overline{B}, \textit{ if } \; B \subset A$$

$$\overline{A} = \overline{B}, \textit{ if } \; B = A$$

- DeMorgan’s Law

$$\overline{(A \cup B)} = \overline{A} \cap \overline{B}$$

$$\overline{(A \cap B)} = \overline{A} \cup \overline{B}$$

Differences

- The difference of two sets, A-B, is the set containing the elements of A that are not elements of B.

$$A - B = A \cap \overline{B} = A - (A \cap B)$$

- Laws for Differences

$$(A - B) \cup B \neq B$$

$$(A \cup A) - A = \varnothing$$

$$(A - A) \cup A = A$$

$$A - \varnothing = A$$

$$A - S = \varnothing$$

$$S - A = \overline{A}$$

Notes and figures are based on or taken from materials in the course textbook: Charles Boncelet, *Probability, Statistics, and Random Signals*, Oxford University Press, February 2016.

## Venn Diagram set theory concepts
## . (2D pictures that can help you understand set theory)

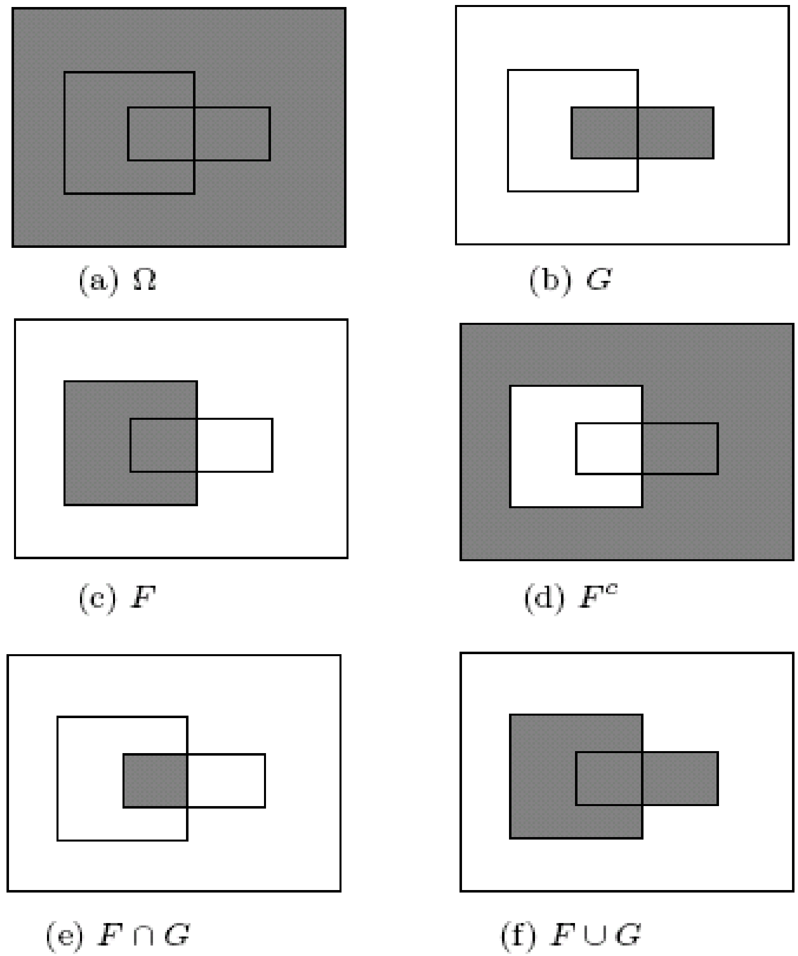



**Figure A.1** Basic set operations

from: Robert M. Gray and Lee D. Davisson, *An Introduction to Statistical Signal Processing*, Cambridge University Press, 2004. Pdf file version found at http://www-ee.stanford.edu/~gray/sp.html

(a) The space
(b) Subset G
(c) Subset F
(d) The Complement of F
(e) Intersection of F and G
(f) Union of F and G

More Venn Diagrams

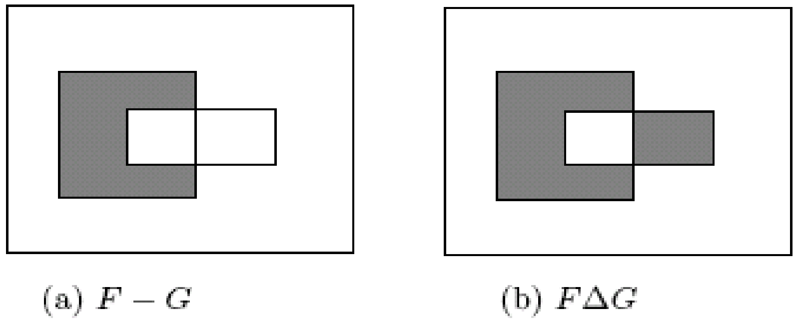



**Figure A.2** Set difference operations

from: Robert M. Gray and Lee D. Davisson, *An Introduction to Statistical Signal Processing*, Cambridge University Press, 2004. Pdf file version found at http://www-ee.stanford.edu/~gray/sp.html

(a) Difference F-G

(b) Difference F-G Union with Difference G-F

$$(F - G) \cup (G - F)$$

If events can be describe in set theory or Venn Diagrams, then probability can directly use the concepts and results of set theory!

What can be said about $\Pr(F \cup G)$? [ read as the probability of event F union event G ]

$$F \cup G = (F - G) \cup G = F \cup (G - F) = F + G - (F \cap G)$$

Therefore,

$$\Pr(F \cup G) = \Pr(F) + \Pr(G) - \Pr(F \cap G)$$

Set algebra is often used to help define probabilities …

## Equalities in Set Algebra

$$F \cup G = G \cup F \text{ commutative law} \tag{A.1}$$
$$F \cup (G \cup H) = (F \cup G) \cup H \text{ associative law} \tag{A.2}$$
$$F \cap (G \cup H) = (F \cap G) \cup (F \cap H) \text{ distributive law} \tag{A.3}$$
$$(F^c)^c = F \tag{A.4}$$
$$F \cap F^c = \emptyset \tag{A.5}$$
$$(F \cap G)^c = F^c \cup G^c \text{ DeMorgan's ``law''} \tag{A.6}$$
$$F \cap \Omega = F \tag{A.7}$$

$$F \cap G = G \cap F \text{ commutative law} \tag{A.8}$$
$$F \cap (G \cap H) = (F \cap G) \cap H \text{ associative law} \tag{A.9}$$
$$(F \cup G)^c = F^c \cap G^c \text{ DeMorgan's other ``law''} \tag{A.10}$$
$$F \cup F^c = \Omega \tag{A.11}$$
$$F \cup \emptyset = F \tag{A.12}$$
$$F \cup (F \cap G) = F = F \cap (F \cup G) \tag{A.13}$$
$$F \cup \Omega = \Omega \tag{A.14}$$
$$F \cap \emptyset = \emptyset \tag{A.15}$$
$$F \cup G = F \cup (F^c \cap G) = F \cup (G - F) \tag{A.16}$$
$$F \cup (G \cap H) = (F \cup G) \cap (F \cup H) \text{ distributive law} \tag{A.17}$$
$$\Omega^c = \emptyset \tag{A.18}$$
$$F \cup F = F \tag{A.19}$$
$$F \cap F = F \tag{A.20}$$

**Table A.1** *Set algebra*

from: Robert M. Gray and Lee D. Davisson, *An Introduction to Statistical Signal Processing*, Cambridge University Press, 2004. Appendix A, Set Theory. Pdf file version found at http://www-ee.stanford.edu/~gray/sp.html

# Axiomatic Definitions Using Sets for Probability

For event A in the "all inclusive" S

$$0 \le \Pr(A) \le 1 \text{ (Axiom 1 \& 2, Theorem 1.3)}$$

$$\Pr(S) = 1 \text{ (Axiom 2)}$$

$$\Pr(\varnothing) = 0 \text{ (Theroem 1.1)}$$

Disjoint Sets

$$\textit{If } A \cap B = \varnothing, \textit{ then } \Pr(A \cup B) = \Pr(A) + \Pr(B)$$

Complement (complementary sets) (defining the complement may be easier sometimes)

$$\textit{If } A \cap \overline{A} = \varnothing, \textit{ then } \Pr(A \cup \overline{A}) = \Pr(A) + \Pr(\overline{A}) = \Pr(S) = 1$$

$$\Pr(A) = 1 - \Pr(\overline{A}) \le 1 \text{ (Theorem 1.2)}$$

Not a Disjoint Sets (solution): (Theorem 1.5)

$$\textit{If } A \cap B \ne \varnothing, \textit{ then } \Pr(A \cup B) = ???$$

- Manipulation (1)

$$A \cup B = A \cup (\overline{A} \cap B), \text{ the union of disjoint sets}$$

$$\Pr(A \cup B) = \Pr[A \cup (\overline{A} \cap B)] = \Pr(A) + \Pr(\overline{A} \cap B)$$

- Manipulation (2)

$$B = (A \cap B) \cup (\overline{A} \cap B), \text{ the union of disjoint sets}$$

$$\Pr(B) = \Pr[(A \cap B) \cup (\overline{A} \cap B)] = \Pr(A \cap B) + \Pr(\overline{A} \cap B)$$

- Manipulation (3)

$$\Pr(\overline{A} \cap B) = \Pr(B) - \Pr(A \cap B), \text{ rearranging from (2)}$$

- Substitution for (1) to produce the desired result

$$\Pr(A \cup B) = \Pr(A) + \Pr(\overline{A} \cap B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$$

Note that we can generally define a bound where

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B) \le \Pr(A) + \Pr(B)$$

equality holds for A and B being disjoint sets!

Notes and figures are based on or taken from materials in the course textbook: Charles Boncelet, *Probability, Statistics, and Random Signals*, Oxford University Press, February 2016.

## Example: 6-sided die

$$\Pr(\alpha_i) = \frac{1}{6}$$

A: The probability of rolling a 1 or a 3, event $A = \{1,3\} = \{1 \cup 3\}$

$$\Pr(A) = \Pr(1) + \Pr(3) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

B: The probability of rolling a 3 or 5, event $B = \{3,5\} = \{3 \cup 5\}$

$$\Pr(B) = \Pr(3) + \Pr(5) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

C: The probability of event A or event B, event $C = \{A \cup B\}$

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

$$\Pr(C) = \frac{1}{3} + \frac{1}{3} - \frac{1}{6} = \frac{1}{2}$$

Note: $C = \{A \cup B\} = \{1,3,5\}$

$$\Pr(C) = \Pr(1) + \Pr(3) + \Pr(5) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$$

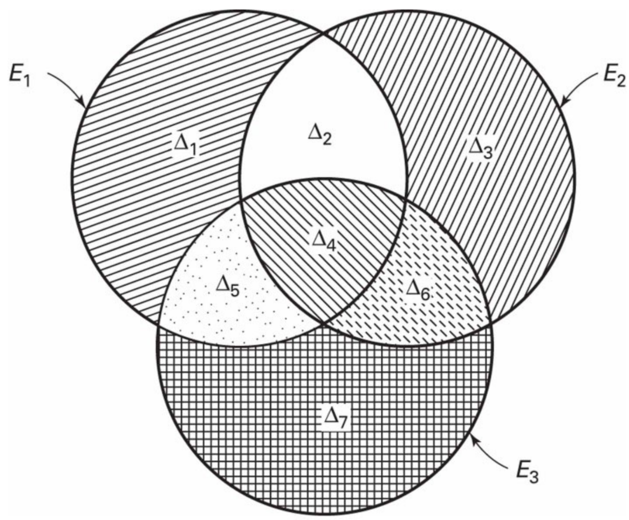
When in doubt, write it out to double check your results!

A Venn diagram may also help.

Notes and figures are based on or taken from materials in the course textbook: Charles Boncelet, *Probability, Statistics, and Random Signals*, Oxford University Press, February 2016.

## Probability of A union of events
## (Theorem 1.6: Inclusion-Exclusion Formula)

An extension of the set theory for unions ….






Probability, Statistics and Random Processes for Engineers, 4th ed., Henry Stark and John W. Woods, Pearson Education, Inc., 2012.

Figure 1.5-1 Partitioning $\bigcup_{i=1}^{3} E_i = \sum_{j=1}^{7} \Delta_j$ into seven disjoint regions Δ1,...,Δ7.)

If $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$, what about $\Pr(A \cup B \cup C) = ???$

$$\begin{aligned}\Pr(A \cup B \cup C) = &\Pr(A) + \Pr(B) + \Pr(C) \\ &- \Pr(A \cap B) - \Pr(A \cap C) - \Pr(B \cap C) \\ &+ \Pr(A \cap B \cap C)\end{aligned}$$

$$\begin{aligned}\Pr(E_1 \cup E_2 \cup E_3) = &\sum_{i=1}^{3} \Pr(E_i) \\ &- \sum_{i=1}^{2} \sum_{j=i+1}^{3} \Pr(E_i \cap E_j) \\ &+ \sum_{i=1}^{1} \sum_{j=i+1}^{2} \sum_{k=j+1}^{3} \Pr(E_i \cap E_j \cap E_k)\end{aligned}$$

Can you recognize a pattern …

- "+"singles … "-"doubles … "+"triples … "-"quads … etc

What about $\Pr(A \cup B \cup C \cup D \cup E \cup F) = ???$

Notes and figures are based on or taken from materials in the course textbook: Charles Boncelet, *Probability, Statistics, and Random Signals*, Oxford University Press, February 2016.

# More Definitions

Probability, the *relative frequency method*:

The number of trials and the number of times an event occurs can be described as

$$N = N_A + N_B + N_C + \cdots$$

the relative frequency is then

$$r(A) = \frac{N_A}{N}$$

note that

$$\frac{N}{N} = \frac{N_A + N_B + N_C + \cdots}{N} = r(A) + r(B) + r(C) + \cdots = 1$$

When experimental results appear with "statistical regularity", the relative frequency tends to approach the probability of the event.

$$\Pr(A) = \lim_{N \to \infty} r(A)$$

and

$$\Pr(A) + \Pr(B) + \Pr(C) + \cdots = 1$$

Where $\Pr(A)$ is defined as the probability of event A.

Mathematical definition of probability:

1. $0 \le \Pr(A) \le 1$
2. $\Pr(A) + \Pr(B) + \Pr(C) + \cdots = 1$, for mutually exclusive events
3. An impossible event, A, can be represented as $\Pr(A) = 0$.
4. A certain event, A, can be represented as $\Pr(A) = 1$.

Odds or probabilities can be assigned to every possible outcome of a "future" trial, experiment, contests, game that has some prior historical basis of events or outcomes.

Notes and figures are based on or taken from materials in the course textbook: Charles Boncelet, *Probability, Statistics, and Random Signals*, Oxford University Press, February 2016.

# Probability Formalized

Axiom 1:

Probability defined between 0 and 1 such that

$$\Pr(A) \geq 0$$

Axiom 2:

Probability for all events in the sample space is 1.0

$$\Pr(S) = 1$$

Axiom 3:

Probability of disjoint sets

$$S = A_1 \cup A_2 \cup A_3 \cdots \cup A_n$$

and

$$A_i \cap A_j = \emptyset, \, for \, i \neq j$$

$$Pr(S) = Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} Pr(A_i)$$

*and total probability would say*

$$Pr(S) = Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} Pr(A_i) = 1.0$$

Where $\Pr(A)$ is defined as the probability of event A.

Notes and figures are based on or taken from materials in the course textbook: Charles Boncelet, *Probability, Statistics, and Random Signals*, Oxford University Press, February 2016.

# Joint Probability – Compound Experiments

Defining probability based on multiple events … two classes for considerations.

- Independent experiments: The outcome of one experiment is not affected by past or future experiments.
    - flipping coins
    - repeating an experiment after initial conditions have been restored
    - Note: these problems are typically easier to solve

- Dependent experiments: The result of each subsequent experiment is affected by the results of previous experiments.
    - drawing cards from a deck of cards
    - drawing straws
    - selecting names from a hat
    - for each subsequent experiment, the previous results change the possible outcomes for the next event.
    - Note: these problems can be very difficult to solve (the "next experiment" changes based on previous outcomes!)

# Independence

Two events, A and B, are independent if and only if

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

Independence is typically assumed when there is no apparent physical mechanism by which the two events could depend on each other. For events derived from independent elemental events, their independence may not be obvious but may be able to be derived.

Independence can be extended to more than two events, for example three, A, B, and C. The conditions for independence of three events is

$$\Pr(A \cap B) = \Pr(A) \cdot \Pr(B) \qquad \Pr(B \cap C) = \Pr(B) \cdot \Pr(C) \qquad \Pr(A \cap C) = \Pr(A) \cdot \Pr(C)$$

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

Note that it is not sufficient to establish pair-wise independence; the entire set of equations is required.

For multiple events, every set of events from n down must be verified. This implies that $2^n - (n+1)$ equations must be verified for n independent events.

## Important Properties of Independence

Unions – help in simplifying the intersection term – if events are independent!

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$$

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A) \cdot \Pr(B)$$

Independent intersection with a Union

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

There will be some example problems where you must determine if events are independent in order to solve the problem.

- switch problems in homework and skills

# Total Probability

For a space, S, that consists of multiple mutually exclusive events, the probability of a random event, B, occurring in space S, can be described based on the conditional probabilities associated with each of the possible events.

Proof:

$$S = A_1 \cup A_2 \cup A_3 \cdots \cup A_n$$

and

$$A_i \cap A_j = \varnothing, \textit{ for } i \neq j$$

$$B = B \cap S = B \cap \left(A_1 \cup A_2 \cup A_3 \cdots \cup A_n\right) = \left(B \cap A_1\right) \cup \left(B \cap A_2\right) \cup \left(B \cap A_3\right) \cdots \cup \left(B \cap A_n\right)$$

$$\Pr(B) = \Pr(B \cap A_1) + \Pr(B \cap A_2) + \Pr(B \cap A_3) \cdots + \Pr(B \cap A_n)$$

But (conditional probability in Chap. 2 – but can be worked out using Venn diagram)

$$\Pr(B \cap A_i) = \Pr(B \mid A_i) \cdot \Pr(A_i), \text{ for } \Pr(A_i) > 0$$

Therefore

$$\Pr(B) = \Pr(B \mid A_1) \cdot \Pr(A_1) + \Pr(B \mid A_2) \cdot \Pr(A_2) + \cdots + \Pr(B \mid A_n) \cdot \Pr(A_n)$$

Remember your math properties: distributive, associative, commutative etc. applied to set theory.

# 1.10 Example: Can S Communicate with D? (Switch problems)

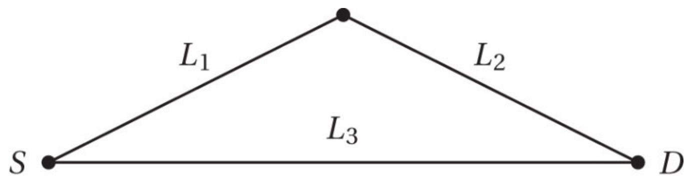



Figure 1.4: Two-path Network

Three links.

If Link 3 is functional, communications occurs.

If both Link 1 and Link 2 are functional communication occurs.

If there are probabilities or likelihoods that Link 1, 2, and 3 are operation, what is the probability you can communicate?

**TABLE 1.1** Start of table of outcomes.

| $L_1$ | $L_2$ | $L_3$ |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 0 | 1 |
| 0 | 1 | 0 |
| 0 | 1 | 1 |
| 1 | 0 | 0 |
| 1 | 0 | 1 |
| 1 | 1 | 0 |
| 1 | 1 | 1 |

Note:

$$Comm = L_3 \cup \left(L_1 \cap L_2\right), \text{ for all links independent!}$$

Notes and figures are based on or taken from materials in the course textbook: Charles Boncelet, *Probability, Statistics, and Random Signals*, Oxford University Press, February 2016.

Consider three independent links and there probabilities

TABLE 1.2 Table of outcomes with probabilities.

| $L_1$ | $L_2$ | $L_3$ | $\Pr[L_1,L_2,L_3]$ |
|---|---|---|---|
| 0 | 0 | 0 | $(1-p)^3$ |
| 0 | 0 | 1 | $p(1-p)^2$ |
| 0 | 1 | 0 | $p(1-p)^2$ |
| 0 | 1 | 1 | $p^2(1-p)$ |
| 1 | 0 | 0 | $p(1-p)^2$ |
| 1 | 0 | 1 | $p^2(1-p)$ |
| 1 | 1 | 0 | $p^2(1-p)$ |
| 1 | 1 | 1 | $p^3$ |

Basic probability

$$\Pr[000] = \Pr[L_1 = 0 \cap L_2 = 0 \cap L_3 = 0]$$

$$\Pr[000] = \Pr[L_1 = 0] \cdot \Pr[L_2 = 0] \cdot \Pr[L_3 = 0]$$

$$\Pr[000] = (1-p) \cdot (1-p) \cdot (1-p) = (1-p)^3$$

$$\Pr[101] = \Pr[L_1 = 1 \cap L_2 = 0 \cap L_3 = 1],$$

$$\Pr[101] = \Pr[L_1 = 1] \cdot \Pr[L_2 = 0] \cdot \Pr[L_3 = 1]$$

$$\Pr[101] = p \cdot (1-p) \cdot p = p^2 \cdot (1-p) = p^2 - p^3$$

$Comm = [S \to D] = L_3 \cup (L_1 \cap L_2)$, for all links independent!

$$Comm = [S \to D] = \{110,001,011,101,111\}$$

$$\Pr[S \to D] = \Pr[\{110,001,011,101,111\}]$$

$$\Pr[S \to D] = \Pr[110] + \Pr[001] + \Pr[011] + \Pr[101] + \Pr[111]$$

$$\Pr[S \to D] = p + p^2 - p^3$$

Alternate Derivation

Notes and figures are based on or taken from materials in the course textbook: Charles Boncelet, *Probability, Statistics, and Random Signals*, Oxford University Press, February 2016.

$$Comm = [S \to D] = L_3 \cup (L_1 \cap L_2)\text{, for all links independent!}$$

$$\begin{aligned}\Pr[S \to D] &= \Pr[L_3 \cup (L_1 \cap L_2)] \\ &= \Pr[L_3] + \Pr[L_1 \cap L_2] - \Pr[L_3] \cdot \Pr[L_1 \cap L_2] \\ &= \Pr[L_3] + \Pr[L_1] \cdot \Pr[L_2] - \Pr[L_3] \cdot \Pr[L_1] \cdot \Pr[L_2]\end{aligned}$$

$$\begin{aligned}\Pr[S \to D] &= \Pr[L_3] + \Pr[L_1] \cdot \Pr[L_2] - \Pr[L_3] \cdot \Pr[L_1] \cdot \Pr[L_2] \\ &= p + p \cdot p - p \cdot p \cdot p \\ &= p + p^2 - p^3\end{aligned}$$

So do you prefer using union and intersection concepts?

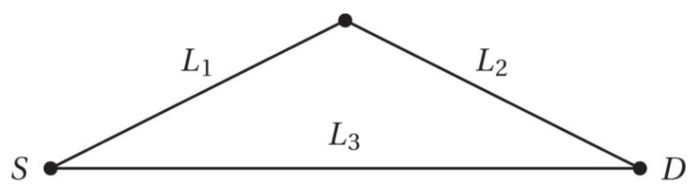



Figure 1.4: Two-path Network

Applying negative logic

$$\Pr\left[\overline{S \to D}\right] = 1 - \Pr[s \to D]$$

$$\Pr\left[\overline{S \to D}\right] = \Pr\left[\overline{L_3} \cap \left(\overline{L_1} \cup \overline{L_2}\right)\right]$$

$$\Pr\left[\overline{S \to D}\right] = \Pr\left[\overline{L_3}\right] \cdot \Pr\left[\left(\overline{L_1} \cup \overline{L_2}\right)\right] = \Pr\left[\overline{L_3}\right] \cdot \left(\Pr\left[\overline{L_1}\right] + \Pr\left[\overline{L_2}\right] - \Pr\left[\overline{L_1}\right] \cdot \Pr\left[\overline{L_2}\right]\right)$$

$$\Pr\left[\overline{S \to D}\right] = \Pr\left[\overline{L_3}\right] \cdot \Pr\left[\left(\overline{L_1} \cup \overline{L_2}\right)\right] = (1-p) \cdot \left((1-p) + (1-p) - (1-p)^2\right)$$

$$\Pr\left[\overline{S \to D}\right] = \Pr\left[\overline{L_3}\right] \cdot \Pr\left[\left(\overline{L_1} \cup \overline{L_2}\right)\right] = (1-p) \cdot \left((2 - 2 \cdot p) - (1 - 2 \cdot p + p^2)\right)$$

$$\Pr\left[\overline{S \to D}\right] = \Pr\left[\overline{L_3}\right] \cdot \Pr\left[\left(\overline{L_1} \cup \overline{L_2}\right)\right] = (1-p) \cdot (1 - p^2) = 1 - p - p^2 + p^3$$

And adjusting for the inverse of the negative result …

$$\Pr[s \to D] = 1 - \Pr\left[\overline{S \to D}\right] = 1 - (1 - p - p^2 + p^3) = p + p^2 - p^3$$

Both negative and positive logic approaches should work!

## A Big Table

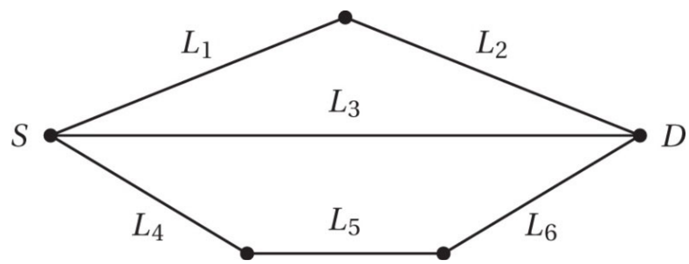
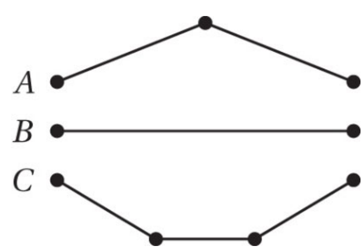



TABLE 1.3 First few rows of a big table.

| $a$ | $L_1$ | $L_2$ | $L_3$ | $L_4$ | $L_5$ | $L_6$ | $\Pr[L_1 \cap L_2 \cap \ldots L_6]$ | $A$ | $B$ | $C$ | $A \cup B \cup C$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $a_0$ | 0 | 0 | 0 | 0 | 0 | 0 | $(1-p)^6$ | 0 | 0 | 0 | 0 |
| $a_1$ | 0 | 0 | 0 | 0 | 0 | 1 | $p(1-p)^5$ | 0 | 0 | 0 | 0 |
| $a_2$ | 0 | 0 | 0 | 0 | 1 | 0 | $p(1-p)^5$ | 0 | 0 | 0 | 0 |
| $a_3$ | 0 | 0 | 0 | 0 | 1 | 1 | $p^2(1-p)^4$ | 0 | 0 | 0 | 0 |
| $a_4$ | 0 | 0 | 0 | 1 | 0 | 0 | $p(1-p)^5$ | 0 | 0 | 0 | 0 |
| $a_5$ | 0 | 0 | 0 | 1 | 0 | 1 | $p^2(1-p)^4$ | 0 | 0 | 0 | 0 |
| $a_6$ | 0 | 0 | 0 | 1 | 1 | 0 | $p^2(1-p)^4$ | 0 | 0 | 0 | 0 |
| $a_7$ | 0 | 0 | 0 | 1 | 1 | 1 | $p^3(1-p)^3$ | 0 | 0 | 1 | 1 |
| $a_8$ | 0 | 0 | 1 | 0 | 0 | 0 | $p(1-p)^5$ | 0 | 1 | 0 | 1 |
| $a_9$ | 0 | 0 | 1 | 0 | 0 | 1 | $p^2(1-p)^4$ | 0 | 1 | 0 | 1 |

Breaking it down into smaller pieces.



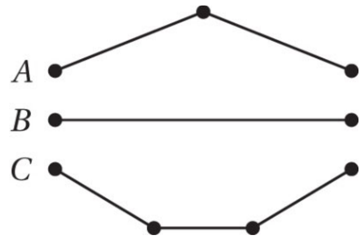

$$Comm = [S \to D] = A \cup B \cup C$$

$$\begin{aligned}\Pr[S \to D] &= \Pr[A \cup B \cup C] \\ &= \Pr[A] + \Pr[B] + \Pr[C] - \Pr[A \cap B] - \Pr[A \cap C] - \Pr[B \cap C] + \Pr[A \cap B \cap C] \\ &= \Pr[A] + \Pr[B] + \Pr[C] - \Pr[A] \cdot \Pr[B] - \Pr[A] \cdot \Pr[C] - \Pr[B] \cdot \Pr[C] + \Pr[A] \cdot \Pr[B] \cdot \Pr[C]\end{aligned}$$

Alternately consider

$$\begin{aligned}\Pr\left[\overline{S \to D}\right] &= \Pr\left[\overline{A} \cap \overline{B} \cap \overline{C}\right] \\ &= \Pr\left[\overline{A}\right] \cdot \Pr\left[\overline{B}\right] \cdot \Pr\left[\overline{C}\right]\end{aligned}$$

Then




$$Comm = [S \to D] = A \cup B \cup C$$

$$Pr[A] = p^2$$

$$Pr[B] = p$$

$$Pr[C] = p^3$$

Alternately consider (negative logic)

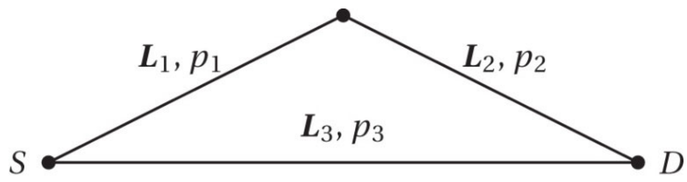
$$\begin{aligned}
\Pr\left[\overline{S \to D}\right] &= \Pr\left[\overline{A} \cap \overline{B} \cap \overline{C}\right] \\
&= \Pr\left[\overline{A}\right] \cdot \Pr\left[\overline{B}\right] \cdot \Pr\left[\overline{C}\right] \\
&= \left(1 - p^2\right) \cdot \left(1 - p\right) \cdot \left(1 - p^3\right) \\
&= \left(1 - p - p^2 + p^3\right) \cdot \left(1 - p^3\right) \\
&= 1 - p - p^2 + p^3 - p^3 + p^4 + p^5 - p^6
\end{aligned}$$

And

$$\begin{aligned}
\Pr[S \to D] &= 1 - \Pr\left[\overline{S \to D}\right] \\
&= 1 - \left(1 - p - p^2 + p^4 + p^5 - p^6\right) \\
&= p + p^2 - p^4 - p^5 + p^6
\end{aligned}$$

Sorry, I like doing things the easiest way …

## Now consider unequal probabilities




Remember the initial derivation?

$$Comm = [S \to D] = L_3 \cup (L_1 \cap L_2), \text{ for all links independent!}$$

$$\begin{aligned}
\Pr[S \to D] &= \Pr[L_3 \cup (L_1 \cap L_2)] \\
&= \Pr[L_3] + \Pr[L_1 \cap L_2] - \Pr[L_3] \cdot \Pr[L_1 \cap L_2] \\
&= \Pr[L_3] + \Pr[L_1] \cdot \Pr[L_2] - \Pr[L_3] \cdot \Pr[L_1] \cdot \Pr[L_2]
\end{aligned}$$

And finally,

$$\begin{aligned}
\Pr[S \to D] &= \Pr[L_3] + \Pr[L_1] \cdot \Pr[L_2] - \Pr[L_3] \cdot \Pr[L_1] \cdot \Pr[L_2] \\
&= p_3 + p_1 \cdot p_2 - p_3 \cdot p_1 \cdot p_2
\end{aligned}$$

Notes and figures are based on or taken from materials in the course textbook: Charles Boncelet, *Probability, Statistics, and Random Signals*, Oxford University Press, February 2016.

## Experiment 1: A bag of marbles, draw 1

A bag of marbles: 3-blue, 2-red, one-yellow

- Objects: Marbles
- Attributes: Color (Blue, Red, Yellow)
- Experiment: Draw one marble, with replacement
- Sample Space: {B, R, Y}
- Probability (relative frequency method)

The probability for each possible event in the sample space is ….

| Event | Probability |
|---|---|
| Blue | 3/6 |
| Red | 2/6 |
| Yellow | 1/6 |
| Total | 6/6 |

This experiment would be easy to run and verify … after lots of trials.

see Matlab Sec1_Marble1.m
ntrials = 6 vs. 600 vs. 6000 (repeat execution a few times)

(Another problem: if we ran 6 trials, what is the probability that we get events that exactly match the probability? 3-Blue, 2-Red, 1 Yellow - a much harder problem)

## Experiment 2: A bag of marbles, draw 2

- Experiment: Draw one marble, replace, draw a second marble.
  "with replacement"

- Sample Space: {BB, BR, BY, RR, RB, RY, YB, YR, YY}

Define the probability of each event in the sample space ….

Joint Probability

- When a desired outcome consists of multiple events. (Read the joint probability of events A and B).

$$\Pr(A,B)$$

Statistically Independent Events

- When the probability of an event does not depend upon any other prior events.
  If trials are performed with replacement and/or the initial conditions are restored, you expect trial outcomes to be independent.

$$\Pr(A,B)=\Pr(B,A)=\Pr(A)\cdot\Pr(B)$$

- The marginal probability of each event is not affected by prior/other events.
  The probability of event A given event B occurred is the same as the probability of event A and vice versa.

$$\Pr(A\mid B)=\Pr(A) \quad \text{and} \quad \Pr(B\mid A)=\Pr(B)$$

- Applicable for multiple objects with single attributes and with replacement.

Therefore

| 1st-rows\2nd-col | 2nd-Blue | 2nd-Red | 2nd-Yellow |
|---|---|---|---|
| 1st-Blue | $\left(\frac{3}{6}\right)\cdot\left(\frac{3}{6}\right)=\frac{9}{36}$ | $\left(\frac{3}{6}\right)\cdot\left(\frac{2}{6}\right)=\frac{6}{36}$ | $\left(\frac{3}{6}\right)\cdot\left(\frac{1}{6}\right)=\frac{3}{36}$ |
| 1st-Red | $\left(\frac{2}{6}\right)\cdot\left(\frac{3}{6}\right)=\frac{6}{36}$ | $\left(\frac{2}{6}\right)\cdot\left(\frac{2}{6}\right)=\frac{4}{36}$ | $\left(\frac{2}{6}\right)\cdot\left(\frac{1}{6}\right)=\frac{2}{36}$ |
| 1-st Yellow | $\left(\frac{1}{6}\right)\cdot\left(\frac{3}{6}\right)=\frac{3}{36}$ | $\left(\frac{1}{6}\right)\cdot\left(\frac{2}{6}\right)=\frac{2}{36}$ | $\left(\frac{1}{6}\right)\cdot\left(\frac{1}{6}\right)=\frac{1}{36}$ |

## Next Concept

***Conditional Probability*** (coming in Chap. 2).

- When the probability of an event depends upon prior events.
  If trials are performed without replacement and/or the initial conditions are not restored, you expect trial outcomes to be dependent on prior results or conditions.

  $$\Pr(A \mid B) \neq \Pr(A) \quad \text{when A follows B}$$

- The joint probability is

  $$\Pr(A,B) = \Pr(B,A) = \Pr(A \mid B) \cdot \Pr(B) = \Pr(B \mid A) \cdot \Pr(A)$$

- Applicable for objects that have multiple attributes and/or for trials performed without replacement.

## Experiment 3: A bag of marbles, draw 2 without replacement

- Experiment: Draw two marbles, without replacement
- Sample Space: {BB, BR, BY, RR, RB, RY, YB, YR}

Therefore

| 1st-rows \ $2^{nd}$-col | $2^{nd}$-Blue | $2^{nd}$-Red | $2^{nd}$-Yellow | $1^{st}$ Marble |
|---|---|---|---|---|
| $1^{st}$-Blue | $\left(\frac{3}{6}\right)\cdot\left(\frac{2}{5}\right)=\frac{6}{30}$ | $\left(\frac{3}{6}\right)\cdot\left(\frac{2}{5}\right)=\frac{6}{30}$ | $\left(\frac{3}{6}\right)\cdot\left(\frac{1}{5}\right)=\frac{3}{30}$ | $\frac{3}{6}$ |
| $1^{st}$-Red | $\left(\frac{2}{6}\right)\cdot\left(\frac{3}{5}\right)=\frac{6}{30}$ | $\left(\frac{2}{6}\right)\cdot\left(\frac{1}{5}\right)=\frac{2}{30}$ | $\left(\frac{2}{6}\right)\cdot\left(\frac{1}{5}\right)=\frac{2}{30}$ | $\frac{2}{6}$ |
| 1-stYellow | $\left(\frac{1}{6}\right)\cdot\left(\frac{3}{5}\right)=\frac{3}{30}$ | $\left(\frac{1}{6}\right)\cdot\left(\frac{2}{5}\right)=\frac{2}{30}$ | $\left(\frac{1}{6}\right)\cdot\left(\frac{0}{5}\right)=\frac{0}{30}$ | $\frac{1}{6}$ |
| $2^{nd}$ Marble | $\frac{3}{6}$ | $\frac{2}{6}$ | $\frac{1}{6}$ | $\frac{6}{6}$ |

## Matlab Marble Simulation Examples:

Sec1_Marble1.m

- example to show small versus large number of sample statistics vs. probability

Sec1_Marble2.m

- example to validate probability and/or small versus large number of trials

Sec1_Marble3.m

- example to validate probability and/or small versus large number of trials