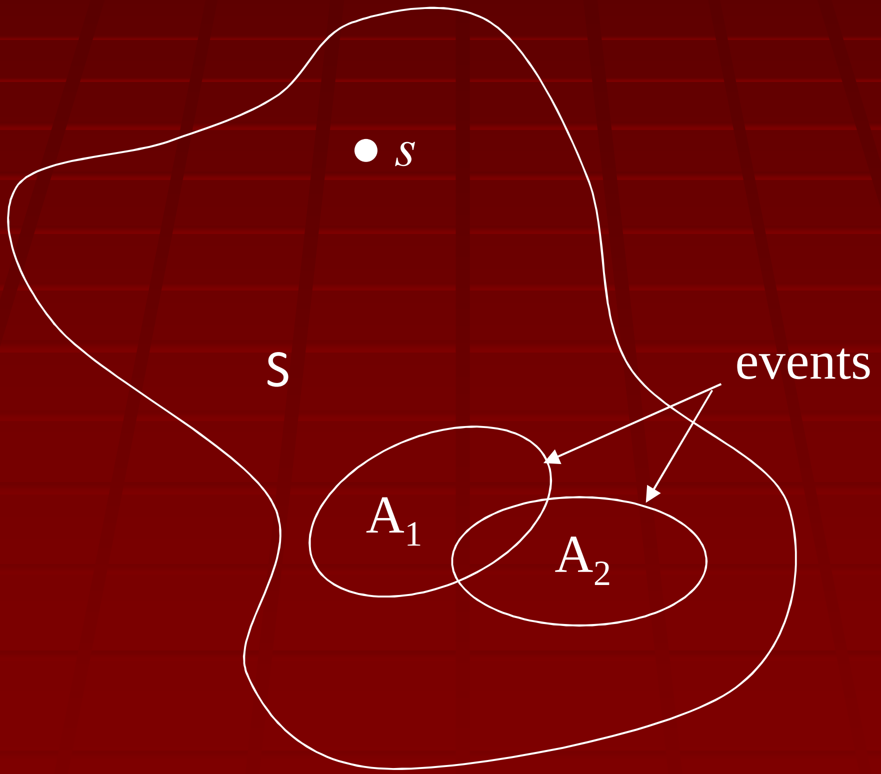
The Probability Model

- Algebra of events
- Probability and conditional probability.
- Applications

  - Network reliability
  - Binary communication
  - Information and coding

Sample Space and Events



Events and Event Operations

S	Certain event
0	Null event
$A_1, A_2, A_3, \ldots$	User defined events
$A_1 + A_2$	Union operation $(A_1 \cup A_2)$
$A_1 A_2$	Intersection operation $(A_1 \cap A_2)$
A^C	Complement operation

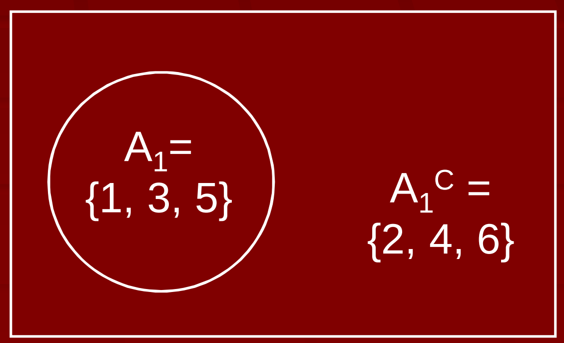
Example:

Consider a sample space represented by $S = \{1,2,3,4,5,6\}$

Let $A_1 = \{1,3,5\}$, $A_2 = \{1,2\}$, and $A_3 = \{2,4,6\}$ be user-defined events.
Find:

- $A_1^C = (\{1,3,5\})^C = \{2,4,6\} = A_3$
- $A_2 + A_3 = \{1,2,4,6\}$
- $A_1 + A_3 = \{1,2,3,4,5,6\} = S$
- $A_1 A_2 = \{1\}$
- $A_1 A_3 = 0$
- $A_1 + 0 = \{1,3,5\} = A_1$
- $A_1\ 0$

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



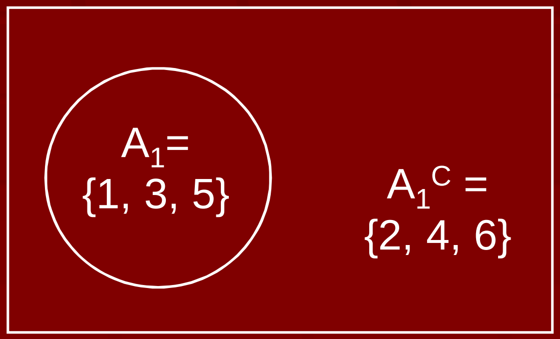
Example:

Consider a sample space represented by $S = \{1,2,3,4,5,6\}$

Let $A_1 = \{1,3,5\}$, $A_2 = \{1,2\}$, and $A_3 = \{2,4,6\}$ be user-defined events.
Find:

- $A_1^C = (\{1,3,5\})^C = \{2,4,6\} = A_3$
- $A_2 + A_3 = \{1,2,4,6\}$
- $A_1 + A_3 = \{1,2,3,4,5,6\} = S$
- $A_1 A_2 = \{1\}$
- $A_1 A_3 = 0$
- $A_1 + 0 = \{1,3,5\} = A_1$
- $A_1 0 = 0$

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



Axioms for the Algebra of Events

- $A_1A_1^C = 0$ Mutual exclusion
- $A_1S = A_1$ Inclusion
- $(A_1^C)^C = A_1$ Double complement
- $A_1 + A_2 = A_2 + A_1$ Commutative law
- $A_1 + (A_2 + A_3) = (A_1 + A_2) + A_3$ Associative law
- $A_1(A_2 + A_3) = A_1A_2 + A_1A_3$ Distributive law
- $(A_1A_2)^C = A_1^C + A_2^C$ De Morgan's law

Additional Identities

- $S^C = 0$
- $A_1 + 0 = A_1$ Inclusion
- $A_1A_2 = A_2A_1$ Commutative law
- $A_1(A_2A_3) = (A_1A_2)A_3$ Associative law
- $A_1 + (A_2A_3) = (A_1 + A_2)(A_1 + A_3)$ Distributive law
- $(A_1 + A_2)^C = A_1^C A_2^C$ De Morgan's law

Finite Unions and Intersections

These are included in the algebra.

$$\bigcup_{i=1}^{N} A_i = A_1 + A_2 + A_3 + \ldots + A_N$$

$$\bigcap_{i=1}^{N} A_i = A_1 A_2 A_3 \ldots A_N$$

Infinite Unions and Intersections

If they are included, the algebra of events is called a Sigma Algebra.

$$\bigcup_{i=1}^{\infty} A_i = A_1 + A_2 + A_3 + \ldots$$

$$\bigcap_{i=1}^{\infty} A_i = A_1 A_2 A_3 \ldots$$

Mutually Exclusive and Collectively Exhaustive Events

Mutually exclusive: $A_k A_j = 0 \qquad k \neq j$

Collectively exhaustive: $\bigcup_j A_j = S$

A set of events that is mutually exclusive and collectively exhaustive is called a *partition*.

Working Definition of the Sample Space

THE SAMPLE SPACE IS REPRESENTED BY THE FINEST GRAIN, MUTUALLY EXCLUSIVE, COLLECTIVELY EXHAUSTIVE SET OF OUTCOMES FOR AN EXPERIMENT.

Discrete Sample Space

Discrete sample space <=> Discrete valued signals

- Output of an 8-bit ADC contains only $2^8 = 256$ values

  $S = \{-5, -4.9609375, \ldots, -0.0390625, 0, 0.0390625, \ldots, 4.9609375\}$

  or $S = \{-128, -127, \ldots, -1, 0, 1, \ldots, 127\}$

  (decimal equivalent of 2's complement representation)

- Discrete sample space has a finite or countably infinite number of outcomes.
  - In this example we have 256 outcomes (finite).

Continuous Sample Space

Continuous sample space <=> Continuous magnitude signals

- Output of the radio receiver is measured at $t = t_1$. The dynamic range of the receiver output is -5V to +5V

$$S = \{s: -5 \le s \le +5\}$$

- A continuous sample space has uncountably infinite values or outcomes
  - *s* could take values like 4.9326784531432677…
- Examples of events and probability assignments:

$$A_1 = \{s: \; -2.5 \le s \le 2.5\}, \quad \Pr[A_1] = 0.50$$

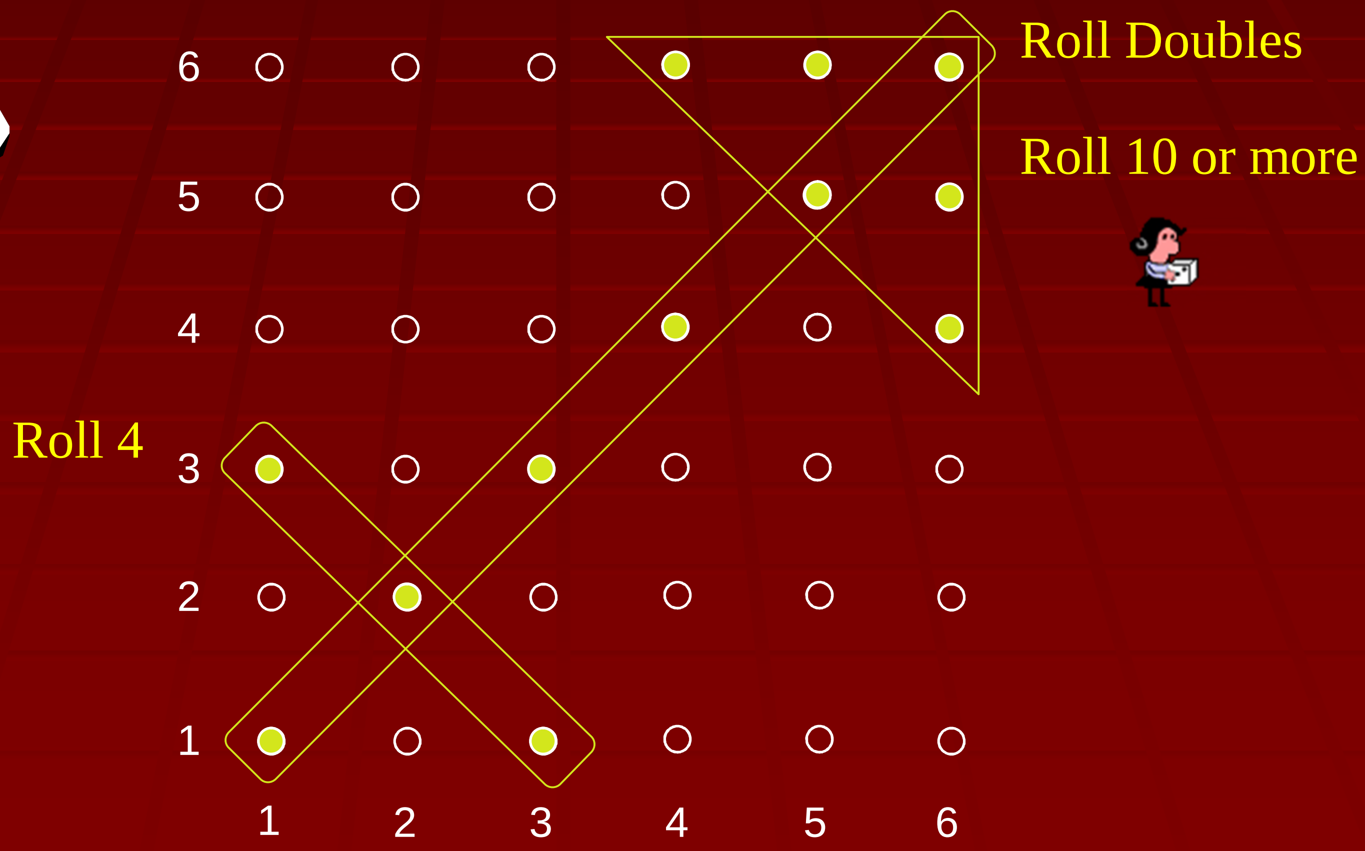
$$A_2 = \{s: \; -1 \le s \le 1\}, \quad \Pr[A_2] = 0.20$$

$$A_3 = \{s: \; s = 2.3 + \Delta x\}, \quad \lim_{\Delta x \to 0} \Pr[A_3] = 0$$

Two Dimensional Sample Space

Roll two dice [discrete sample space]

$S = \{(i,j): (1,1), (1,2), (1,3), \ldots, (4,3), \ldots, (6,4), (6,5), (6,6)\}$



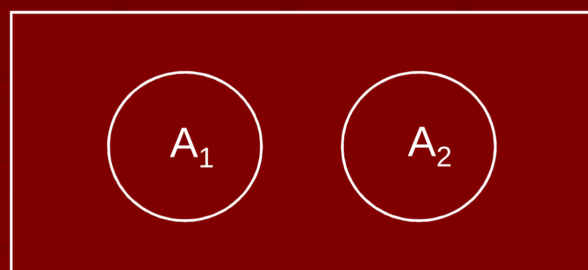
Axioms of Probability

I. $\Pr[A_i] \geq 0$ for any event

II. $\Pr[S] = 1$

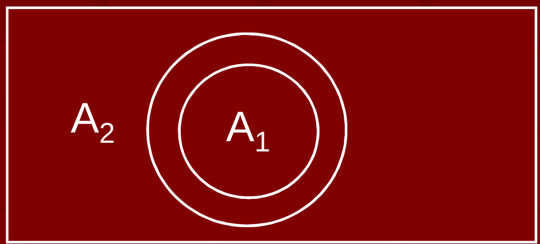
III(a). If $A_1 A_2 = 0$, then $\Pr[A_1 + A_2] = \Pr[A_1] + \Pr[A_2]$

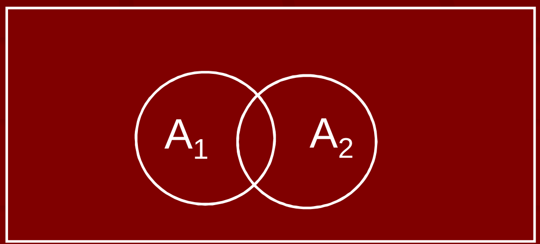
III(b). If $A_i A_j = 0$ for $i \neq j$, then $\Pr\left[\bigcup_{i=1}^{\infty} A_i\right] = \sum_{i=1}^{\infty} \Pr[A_i]$



Some Corollaries

1. $\Pr[A^c] = 1 - \Pr[A]$ (ii, iii)
2. $0 \le \Pr[A] \le 1$ (i, ii, iii)
3. If $A_1 \subset A_2$, then $\Pr[A_1] \le \Pr[A_2]$ (i, iii)
4. $\Pr[0] = 0$ (ii, iii)
5. If $A_1 A_2 = 0$, then $\Pr[A_1 A_2] = 0$
6. $\boxed{\Pr[A_1 + A_2] = \Pr[A_1] + \Pr[A_2] - \Pr[A_1 A_2]}$





The Principle of Total Probability

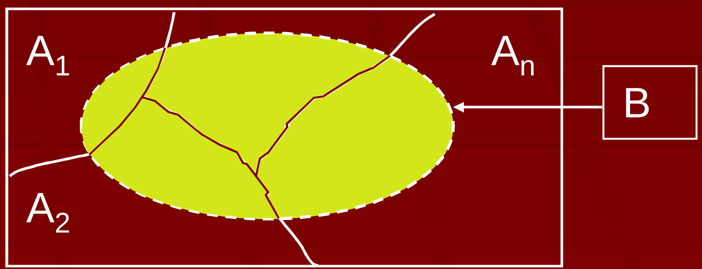
Let $A_1, A_2, \ldots, A_n$ be a set of mutually exclusive and collectively exhaustive events:

$$A_k A_j = 0 \qquad k \neq j$$

$$\bigcup_{j=1}^{n} A_j = S \quad \text{then} \quad \sum_{j=1}^{n} \Pr[A_j] = 1$$

Now let B be any event in S. Then,

$\Pr[B] = \Pr[BA_1] + \Pr[BA_2] + \ldots + \Pr[BA_n]$



Independence of Events

Two events A_1 and A_2 are said to be statistically independent if and only if

$$\Pr[A_1 A_2] = \Pr[A_1]\,\Pr[A_2]$$

When the situation claims that events are “independent,” this is the one and only test.

System Reliability Calculations

Series connection of switches

Assume that switch failures are <u>independent</u>; and failure results in an open connection.

$$F = \{\text{no connection between x and y}\} = A_1 + A_2$$

The probability that the connection fails:

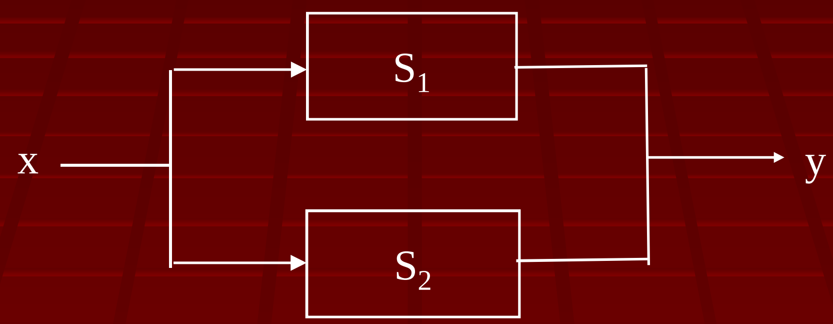
$$\Pr[F] = \Pr[A_1 + A_2] = \Pr[A_1] + \Pr[A_2] - \Pr[A_1 A_2]$$

$$= \Pr[A_1] + \Pr[A_2] - \Pr[A_1]\Pr[A_2] \qquad (A_1 \text{ and } A_2 \text{ are independent})$$

$$= p + p - p^2 = 2p - p^2$$

System Reliability Calculations

Parallel Connection of switches



Define: $A_1 = \{S_1 \text{ fails}\}$, $\Pr[A_1] = p$, $\Pr[A_1^c] = 1 - p = q$

$A_2 = \{S_2 \text{ fails}\}$, $\Pr[A_2] = p$, $\Pr[A_2^c] = 1 - p = q$

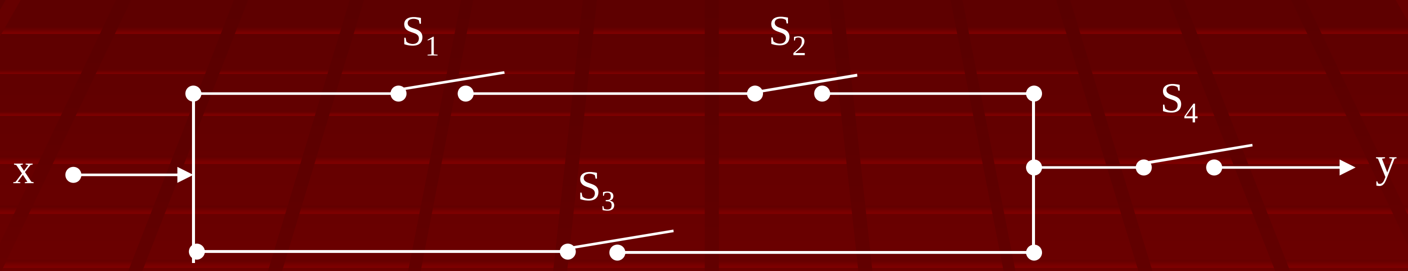
$F = \{\text{no connection between x and y}\} = A_1A_2$

The probability that the connection fails:

$$\Pr[F] = \Pr[A_1A_2] = \Pr[A_1]\Pr[A_2] = p^2 \qquad (A_1 \text{ and } A_2 \text{ are independent})$$

Combination Example

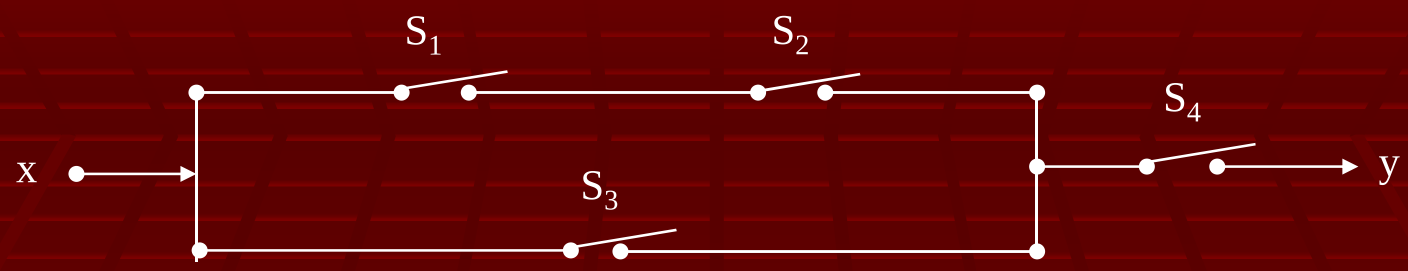
Consider a simple switching network as follows:



Define: A_k = {switch S_k is open}, $k = 1,2,3,4$ $\Pr[A_k] = p$

Let F = {no connection between x and y} $= (A_1 + A_2)A_3 + A_4$

Combination Example (cont'd)



(a) Find the probability that the path between x and y is established.

The probability of path failure is given by

$$\rightarrow \Pr[F] = \Pr[(A_1 + A_2)A_3 + A_4] = \Pr[(A_1A_3 + A_2A_3) + A_4]$$

$$= \Pr[A_1A_3 + A_2A_3] + \Pr[A_4] - \Pr[A_1A_3A_4 + A_2A_3A_4]$$

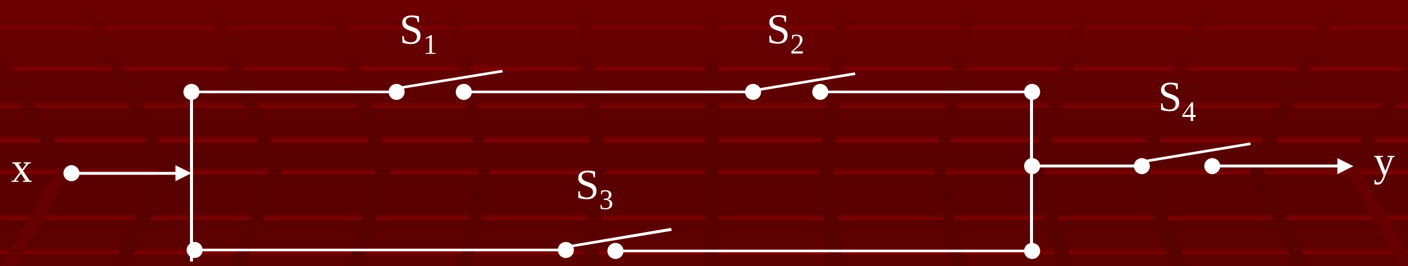
$$= \Pr[A_4] + \Pr[A_1A_3] + \Pr[A_2A_3] - \Pr[A_1A_2A_3]$$

$$- \Pr[A_1A_3A_4] - \Pr[A_2A_3A_4] + \Pr[A_1A_2A_3A_4]$$

$$= p + p^2 + p^2 - p^3 - p^3 - p^3 + p^4$$

$$= p + 2p^2 - 3p^3 + p^4$$

Combination Example (cont'd)



(a) Find the probability that the path between x and y is established.

The probability of path failure is given by

$$\rightarrow Pr[F] = Pr[(A_1 + A_2)A_3 + A_4] = Pr[(A_1A_3 + A_2A_3) + A_4]$$
$$= Pr[A_1A_3 + A_2A_3] + Pr[A_4] - Pr[A_1A_3A_4 + A_2A_3A_4]$$
$$= Pr[A_4] + Pr[A_1A_3] + Pr[A_2A_3] - Pr[A_1A_2A_3]$$
$$- Pr[A_1A_3A_4] - Pr[A_2A_3A_4] + Pr[A_1A_2A_3A_4]$$
$$= p + p^2 + p^2 - p^3 - p^3 - p^3 + p^4$$
$$= p + 2p^2 - 3p^3 + p^4$$

The desired probability is then given by

$$Pr[\text{path established}] = 1 - Pr[F] = 1 - p - 2p^2 + 3p^3 - p^4$$

Probability as a function of parameter ‘p’

$$\Pr[\text{path established}] = 1 - \Pr[F] = 1 - p - 2p^2 + 3p^3 - p^4$$

p	1 - Pr[F]
0.1	0.88290000
0.01	0.98980299
0.001	0.99899800
0.0001	0.99989998

Repeated Independent Trials (Bernoulli Trials)

A random experiment, E, consists of several sub-experiments or trials, E_i:

- All sub-experiments have the same sample space, S_i.
- Events from all sub-experiments are mutually independent:
  $\Pr[A_1 A_2 A_3 \ldots A_n] = \Pr[A_1]\Pr[A_2]\Pr[A_3]\ldots\Pr[A_n]$,
  where A_i is an event from S_i.
- The sample space of E is:
  $S = S_1 \times S_2 \times S_3 \times \ldots \times S_n$

Example of Bernoulli Trials

A sequence of 4 bits is transmitted over a channel. Bit errors occur with probability 0.05. What is the sample space and what are the probabilities of the outcomes?

For a single bit we have 2 sub-events:

Φ	no error	$\Pr[\Phi] = 0.95$
E	error	$\Pr[E] = 0.05$

The sample space and probabilities are:

ΦΦΦΦ	ΦΦΦE	ΦΦEΦ	ΦΦEE	…	EEEE
$(0.95)^4$	$(0.95)^3(0.05)$	$(0.95)^2(0.05)(0.95)$	$(0.95)^2(0.05)^2$	…	$(0.05)^4$
0.8145	0.0429	0.0429	0.0023	…	6.25E-6

Counting Methods and Probabillity

The assignment of probability is given by

$$\Pr[\mathrm{A}_1] = \frac{\text{Number of outcomes in Event } \mathrm{A}_1}{\text{Number of outcomes of experiment}}$$

The rule of products:

Consider an experiment with *n* outcomes; repeat it *r* times

The total number of outcomes is given by

$$n \cdot n \cdot \cdots \cdot n = n^r$$

or in the general case

$$n_1 n_2 \ldots n_r = \prod_{i=1}^{r} n_i$$

Examples

- Roll a die 4 times sequentially. The total number of outcomes is

$$6 \cdot 6 \cdot 6 \cdot 6 = 6^4 = 1296$$

- Form 5-letter words using 26 English alphabet characters. Characters can be repeated, and the words so formed do not have to be meaningful.

$$26^5 = 11{,}881{,}376$$

- Construct variable names of length 3 using a letter, a number, and a letter (e.g., A2C).

$$26 \cdot 10 \cdot 26$$

Permutations

(products without replacement)

Select r objects from among a given set of n distinct objects where we pay attention to the order in which the r objects are selected.

$$P_r^n = n(n-1)(n-2)\Lambda\ (n-r+1)$$

$$= \frac{n!}{(n-r)!}, \qquad \text{for } \ r \leq n$$

Important special case:

For $r = n$: $\quad P_n^n = n!$

Permutation Example

Form 5-letter words using the English alphabet. The characters cannot be repeated, and the words do not have to be meaningful. The total number of words that can be formed is:

$$P_5^{26} = \frac{26!}{(26-5)!}$$

$$= \frac{26!}{21!} = 26 \cdot 25 \cdot 24 \cdot 23 \cdot 22 = 7{,}893{,}600$$

Combinations

Select r objects from among a given set of n distinct objects where we pay no attention to the order in which the r objects are selected.

$$C_r^n = \binom{n}{r} = \frac{n!}{r!(n-r)!} = \frac{n(n-1)\cdots(n-r+1)}{r(r-1)(r-2)\cdots(1)}$$

Binomial coefficient “n *choose* r”

Permutations and Combinations Example

Consider 5 workstations having equal capabilities: {a, b, c, d, e}

- *Permutations*: Select two workstations where one will be a server and the other a graphics workstation. The possible selections are:

  ab ba ac ca ad da ae ea bc cb bd db be eb cd dc ce ec de ed

$$P_2^5 = \frac{5!}{(5-2)!} = 5 \cdot 4 = 20$$

"ab ≠ ba"

- *Combinations*: Select two workstations where both will be used as graphics workstations. The possible selections are:

  ab ac ad ae bc bd be cd ce de

$$\binom{5}{2} = \frac{5!}{2!(5-2)!} = \frac{5 \cdot 4}{2} = 10$$

"ab = ba"

Buying Computers

We plan to buy 5 personal computers. The computer store has a stock of 10 foreign-made PCs and 15 US-made PCs that meet our specifications.

(a) Assuming that the 5 computers are randomly chosen from this lot, what is the probability that exactly 3 US-made computers are selected?

- The sample space is given by

  S = {combinations of r = 5 chosen from n = 25}

$$N_S = \binom{25}{5} = \frac{25!}{5!(25-5)!}$$

- The desired event A = {*exactly* 3 of the 5 selected are US-made}

$$N_A = \binom{15}{3}\binom{10}{2} = \frac{15!\,10!}{3!(15-3)!\;\;2!(10-2)!}$$

Buying Computers (cont'd)

- The probability that we have 3 US-made computers is

$$\Pr[A] = \frac{N_A}{N_S} = \frac{\binom{15}{3}\binom{10}{2}}{\binom{25}{5}} = \frac{\left(\frac{15!}{3!\ (15-3)!}\right)\left(\frac{10!}{2!\ (10-2)!}\right)}{\frac{25!}{5!\ (25-5)!}}$$

$$= \frac{15!\ \ 10!\ \ 5!\ \ (25-5)!}{3!\ \ (15-3)!\ \ 2!\ \ (10-2)!\ \ 25!} \cong 0.3854$$

[Hyper geometric distribution]

Buying Computers (cont'd)

(b) Assuming that the 5 computers are randomly chosen from this lot, what is the probability that *at least* 1 is foreign made?

- The desired event B = {one or more of the 5 are foreign made}
- Let event C = {none of the 5 selected computers is foreign-made}
- Since B = C^C, we can write Pr[B] = 1 - Pr[C]

$$\Pr[C] = \frac{N_B}{N_S} = \frac{\binom{15}{5}\binom{10}{0}}{\binom{25}{5}} = \frac{15! \quad 10! \quad 5! \quad (25-5)!}{5! \quad (15-5)! \quad 0! \quad (10-0)! \quad 25!} \cong 0.0565$$

$$\Pr[B] = 1 - \Pr[C] = 1 - 0.0565 = 0.9435$$

Another Example

Consider a box of 25 modem chips; 5 of them are known to be defective.

Select 6 from the box at random and test them. What is the probability that exactly 2 are defective?

- Sample Space: S = {combinations of r = 6 chosen from n = 25}
- Event: A = {*exactly* 2 of the 6 selected chips are defective}
- The number of outcomes in S is given by:

$$N_S = \binom{25}{6} = \frac{25!}{6!(25-6)!}$$

- For the selected 6 chips, we are interested in the case, where 2 are defective (i.e., they are from the 5 defective chips in the box) and 4 are non-defective (i.e., they are from the 20 non-defective chips in the box).

Example (cont'd)

- The number of outcomes in A is given by:

$$N_A = \binom{5}{2}\binom{20}{4} = \frac{5!\ 20!}{2!(5-2)!\ 4!(20-4)!}$$

- The probability that exactly 2 of the 6 selected chips are defective is

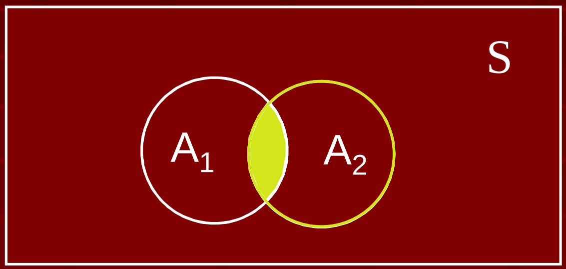
$$\Pr[A] = \frac{N_A}{N_S} = \frac{\binom{5}{2}\binom{20}{4}}{\binom{25}{6}} = \frac{5!\quad 20!\quad 6!\quad (25-6)!}{2!\quad (5-2)!\quad 4!\quad (20-4)!\quad 25!}$$

$$\cong 0.2736$$

Conditional Probability

Probability of occurrence of one event (say, A_1) subject to the knowledge that another event (say, A_2) has occurred.

$$\Pr\left[A_1 \middle| A_2\right] = \frac{\Pr[A_1 A_2]}{\Pr[A_2]}$$



$\Pr[A_1 \mid A_2]$ is read as "probability of A_1 *given* A_2 "

If A_1 and A_2 are independent, then

$$\Pr\left[A_1 \middle| A_2\right] = \frac{\Pr[A_1 A_2]}{\Pr[A_2]} = \frac{\Pr[A_1]\Pr[A_2]}{\Pr[A_2]} = \Pr[A_1]$$

Conditional Probability Example

Consider a sequence of 3 binary numbers (occurring randomly).

Sample space: $S = \{000, 001, 010, 011, 100, 101, 110, 111\}$

Find the probability of more 1's than 0's given the first bit is a 1.

- Define two events:

  $A_1 = \{\text{more 1's than 0's}\} = \{011, 101, 110, 111\}$ (0.5)

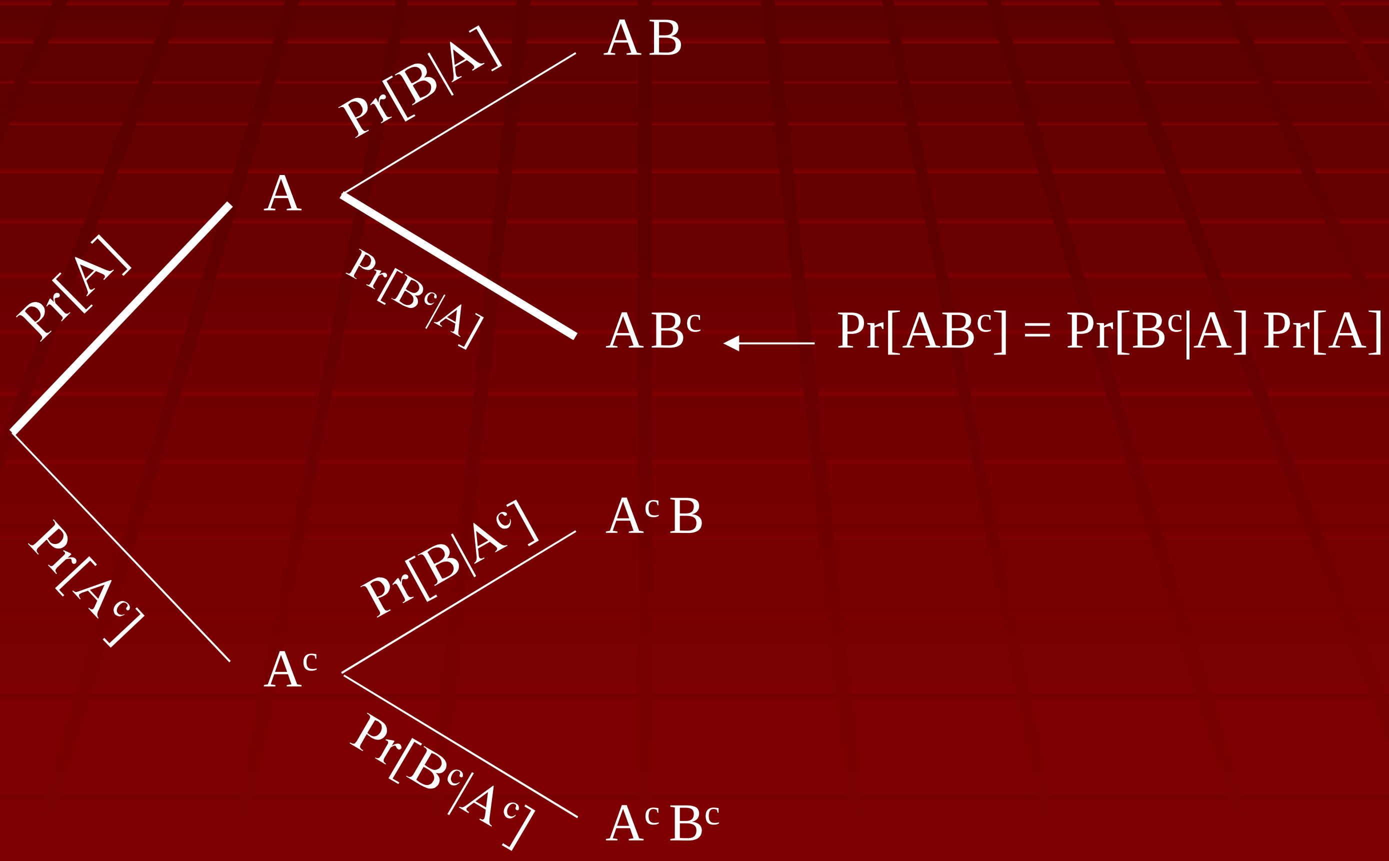
  $A_2 = \{\text{the first bit is a 1}\} = \{100, 101, 110, 111\}$ (0.5)

- Their intersection:

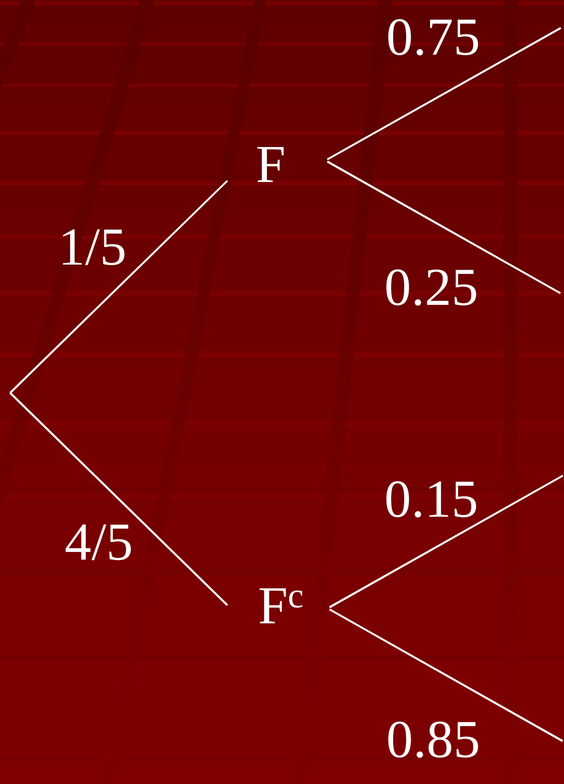
  $A_1A_2 = \{101, 110, 111\}$ (0.375)

$$\Pr[A_1|A_2] = \frac{\Pr[A_1A_2]}{\Pr[A_2]} = \frac{0.375}{0.5} = 0.75$$

Event Tree Used to Form a Sample Space



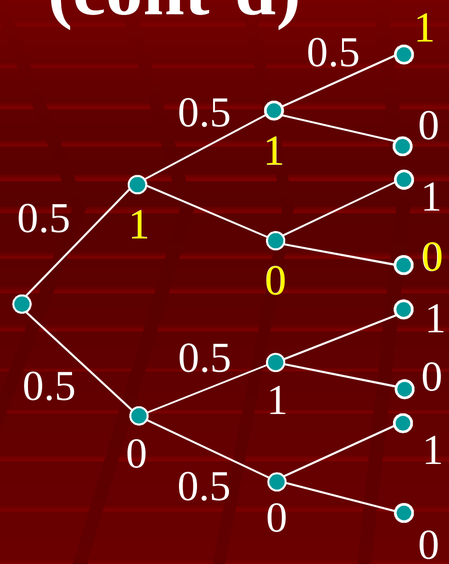
Computation of Probabilities



Sample Space	Probability
$F\,R$	(1/5)(0.75)
$F\,R^c$	(1/5)(0.25)
$F^c\,R$	(4/5)(0.15)
$F^c\,R^c$	(4/5)(0.85)

Example (cont'd)

Tree diagram



Sample Space	Probability
111	0.125
110	0.125
101	0.125
100	0.125
011	0.125
010	0.125
001	0.125
000	0.125

Event A_1 = {more 1's than 0's}

Event A_2 = {the first bit is a 1}

A_1A_2

All 8 events in the sample space have probability 1/8, therefore

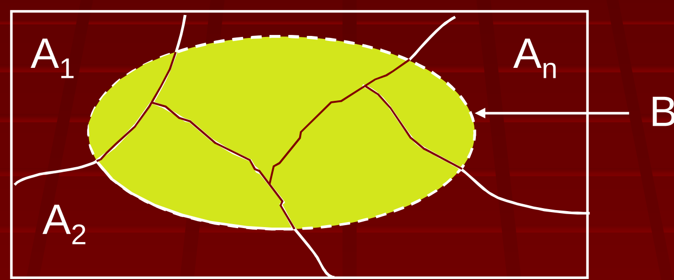
$$\Pr[A_2] = \frac{4}{8} \quad \text{and} \quad \Pr[A_1A_2] = \frac{3}{8}$$

The conditional probability is obtained as follows:

$$\Pr\left[A_1 \middle| A_2\right] = \frac{\Pr[A_1A_2]}{\Pr[A_2]} = \frac{3/8}{4/8} = \frac{3}{4}$$

Principle of Total Probability Revisited

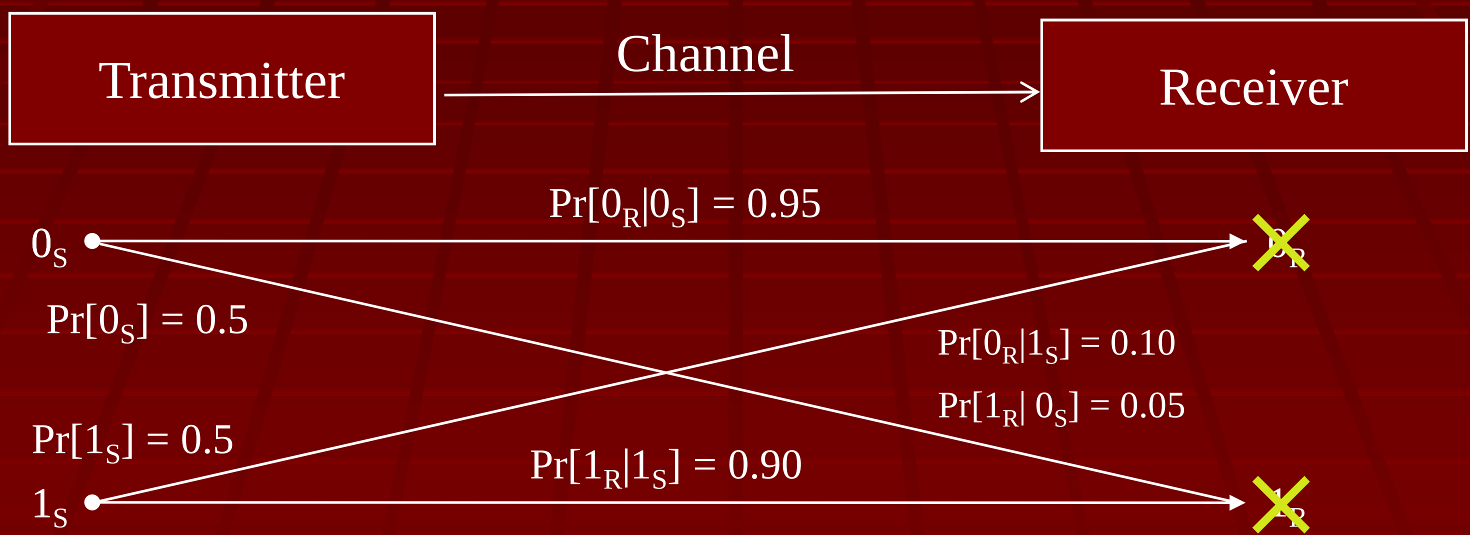
Let $A_1, A_2, \ldots, A_n$ be mutually exclusive and collectively exhaustive events. Let B be an event in S.



Then,

$$
\begin{aligned}
\Pr[B] &= \Pr[BA_1] + \Pr[BA_2] + \ldots + \Pr[BA_n] \\
&= \Pr\left[B \middle| A_1\right]\Pr[A_1] + \ldots + \Pr\left[B \middle| A_n\right]\Pr[A_n] \\
&= \sum_{i=1}^{n} \Pr\left[B \middle| A_i\right]\Pr[A_i]
\end{aligned}
$$

Binary Communication Channel



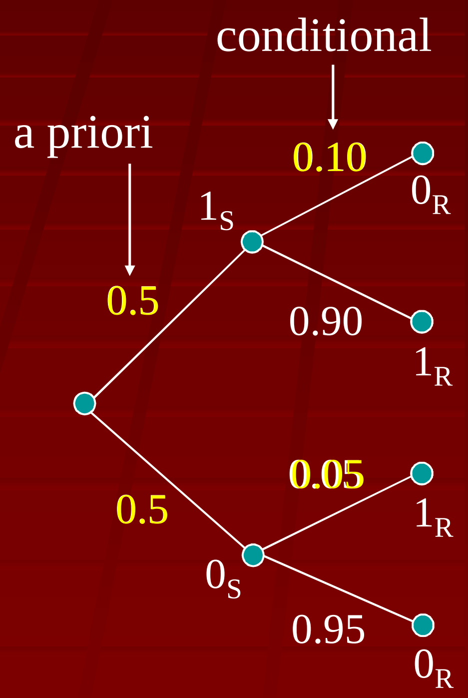
We can compute the probability of joint events:

$Pr[0_S 0_R] = Pr[0_R|0_S]Pr[0_S] = (0.95)(0.5) = 0.475$

$Pr[0_S 1_R] = Pr[1_R|0_S]Pr[0_S] = (0.05)(0.5) = 0.025$ (error)

Binary Communication Channel (cont'd)

Computing the probability of error:

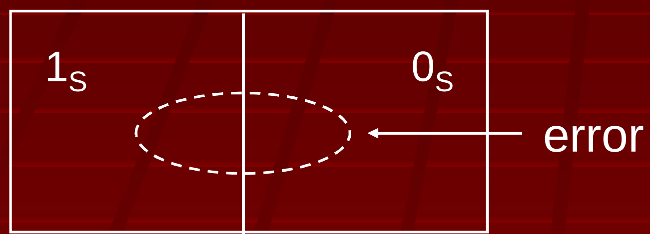


	Probability
$1_S 0_R$ (error)	0.050
$1_S 1_R$	0.450
$0_S 1_R$ (error)	0.025
$0_S 0_R$	0.475

$$\Pr[\text{error}] = 0.050 + 0.025 = 0.075$$

Binary Communication Channel (cont'd)

Compute the probability of error another way:



$$\Pr[\text{error} \mid 1_S] = \Pr[0_R \mid 1_S] = 0.10$$
$$\Pr[\text{error} \mid 0_S] = \Pr[1_R \mid 0_S] = 0.05$$

Use the Law of Total Probability:

$$\begin{aligned}\Pr[\text{error}] &= \Pr[\text{error} \mid 1_S]\Pr[1_S] + \Pr[\text{error} \mid 0_S]\Pr[0_S] \\ &= 0.10 \cdot 0.50 + 0.05 \cdot 0.50 = 0.075\end{aligned}$$

More on Conditional Probability

From the definition of conditional probability, we can write

$$\Pr\left[A_1 \middle| A_2\right] = \frac{\Pr[A_1 A_2]}{\Pr[A_2]} \quad \text{or} \quad \Pr[A_1 A_2] = \Pr\left[A_1 \middle| A_2\right]\Pr[A_2]$$

$$\Pr\left[A_2 \middle| A_1\right] = \frac{\Pr[A_1 A_2]}{\Pr[A_1]} \quad \text{or} \quad \Pr[A_1 A_2] = \Pr\left[A_2 \middle| A_1\right]\Pr[A_1]$$

$$\therefore \quad \Pr\left[A_1 \middle| A_2\right] \Pr[A_2] = \Pr\left[A_2 \middle| A_1\right]\Pr[A_1]$$

This can be written as

$$\boxed{\Pr\left[A_2 \middle| A_1\right] = \frac{\Pr\left[A_1 \middle| A_2\right]\Pr[A_2]}{\Pr[A_1]}}$$

Bayes' Rule

Let $A_1, A_2, \ldots, A_n$ be a set of mutually exclusive, collective exhaustive events (partition). Then,

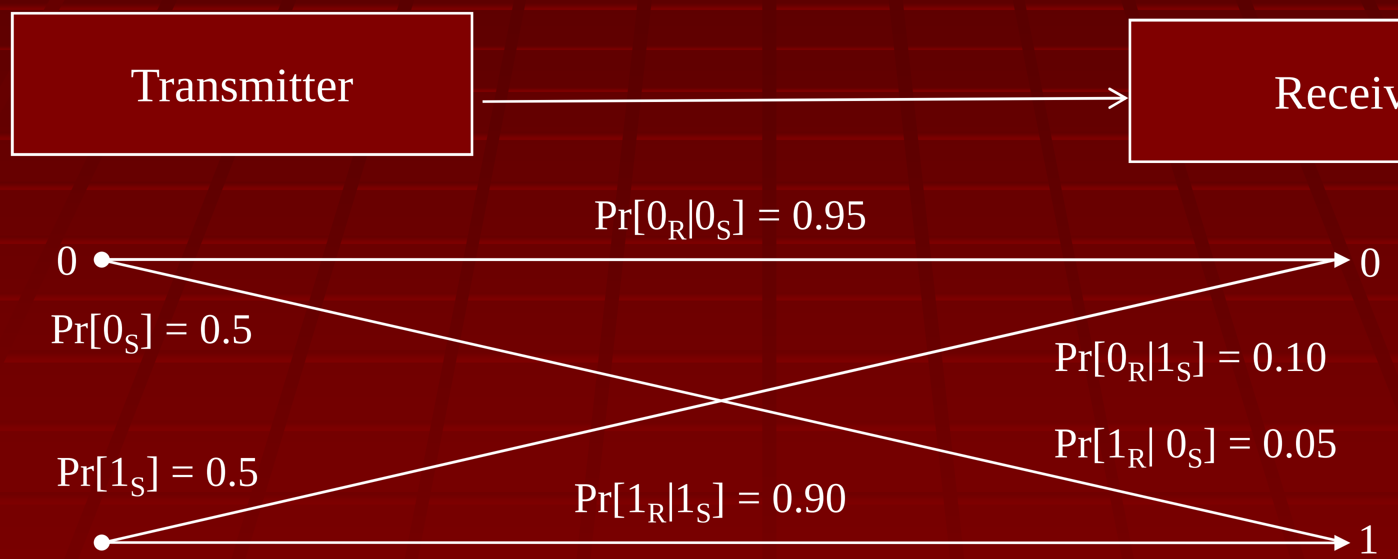
$$\Pr\left[A_j \mid B\right] = \frac{\Pr[B \mid A_j]\Pr[A_j]}{\Pr[B]}$$

Or, applying the Principle of Total Probability

$$\Pr\left[A_j \mid B\right] = \frac{\Pr[B \mid A_j]\Pr[A_j]}{\sum_{k=1}^{n} \Pr[B \mid A_k]\Pr[A_k]}$$

This is called Bayes' Rule.

Binary Communication Channel Revisited



Determine the inverse probability, $P[1_S\ |1_R]$.

Binary Communication Channel (cont'd)

Determine the inverse probability, $P[1_S \mid 1_R]$:

$$\Pr\left[1_S \mid 1_R\right] = \frac{\Pr[1_R \mid 1_S]\Pr[1_S]}{\Pr[1_R]}$$

$$= \frac{\Pr[1_R \mid 1_S]\Pr[1_S]}{\Pr[1_R \mid 1_S]\Pr[1_S] + \Pr[1_R \mid 0_S]\Pr[0_S]}$$

$$= \frac{0.45}{0.45 + 0.025} = 0.9474$$

Integrated Circuit Example

Consider 3 boxes of integrated circuits (ICs).

Box 1 contains 1500 ICs and 10% of them are defective;

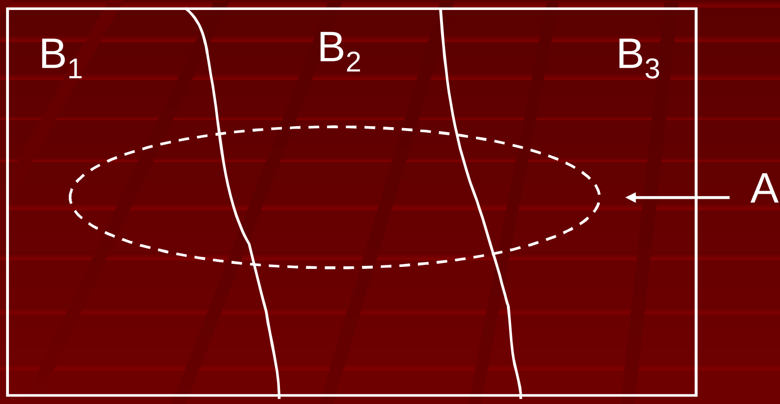
Box 2 contains 2000 ICs and 20% of them are defective; and

Box 3 contains 3000 ICs and 16% of them are defective.

Select 1 of the 3 boxes at random and choose an IC from that box at random.

Integrated Circuit Example (cont'd)

(a) What is the probability that this IC is defective?



Define: A = "selected IC is defective" ,

B_i = "IC is from box i"

By the principle of total probability, we can write

$$\Pr[A] = \Pr\left[A \middle| B_1\right]\Pr[B_1] + \Pr\left[A \middle| B_2\right]\Pr[B_2] + \Pr\left[A \middle| B_3\right]\Pr[B_3]$$

$$= 0.10 \cdot \frac{1}{3} + 0.20 \cdot \frac{1}{3} + 0.16 \cdot \frac{1}{3} = \frac{0.46}{3} = 0.1533$$

Integrated Circuit Example (cont'd)

(b) Suppose that the selected IC is found to be defective. What is the probability that this IC came from box #3?

By Bayes' theorem, we can write

$$\Pr\left[B_3 \middle| A\right] = \frac{\Pr[A \mid B_3]\Pr[B_3]}{\Pr[A]}$$

$$= \frac{0.16 \cdot \dfrac{1}{3}}{0.1533} = \frac{0.0533}{0.1533} = 0.3479$$

(c) Suppose all IC's are mixed in one box and an IC is selected at random. What is the probability that the IC is defective?

Basic Information Theory

Given an event A and its probability Pr[A], the “information” associated with A is defined by

$$I[\mathrm{A}] = \log_x \frac{1}{\Pr[\mathrm{A}]} = -\log_x \Pr[\mathrm{A}],$$

where x is the base of the logarithm:

if $x = 2$, the units of information are bits;

if $x = 10$, the units are hartleys;

and if $x = e$, the units are nats.

Note the identity: $\log_a b = x$ means that $a^x = b$.

Information Theory Example

Consider two events: A_1 and A_2 with corresponding probabilities of occurrence of 0.125 and 0.875, respectively.

The information associated with these events:

$$I[A_1] = -\log_2 (0.125) = 3 \text{ bits}$$

$$I[A_2] = -\log_2 (0.875) = 0.1925 \text{ bits}$$

Entropy

Given a set of independent events that are mutually exclusive and collectively exhaustive, the average information associated with the random experiment is defined as

$$H = \sum_i \Pr[A_i] \cdot I[A_i] = -\sum_i \Pr[A_i] \cdot \log_x \Pr[A_i]$$

Entropy Example

Consider a sequence 1 2 3 2 3 4 5 4 5 6 7 8 9 8 9 0.

We estimate the probability of occurrence of each symbol as follows:

Pr[1] = Pr[6] = Pr[7] = Pr[0] = 1/16

Pr[2] = Pr[3] = Pr[4] = Pr[5] = Pr[8] = Pr[9] = 2/16

The entropy of this sequence is

$$H = -\sum_i \Pr[A_i] \cdot \log_2 \Pr[A_i] = 3.25 \text{ bits}$$

Shannon-Fano Code

Messages are composed of an alphabet in which the frequency of occurrence of each letter is a probabilistic phenomenon.

- For transmission purposes the messages are compressed such that the code length of a letter is inversely proportional to its frequency of occurrence (e.g., think of the Morse code).
- Since the letters are transmitted sequentially, no short codeword can be part of the start of a longer codeword for unique decodability.

Shannon-Fano Algorithm:

- Arrange letters in a descending order of their probabilities by breaking any ties arbitrarily.
- Starting at the top, partition the letters into two equi-probable subgroups (as closely as possible): assign 0 to the first subgroup and 1 to the second.
- Continue partitioning the subgroups until all letters are exhausted: after each partition, assign a 0 to the first group and a 1 to the second and append the newly assigned bits to the previously assigned bits.

Example

Given a text message "ELECTRICAL ENGINEERING," determine the relative probabilities of the letters in the message and find the Shannon-Fano code for each letter. Ignore the space character.

- Since there are 21 letters in the message, we have the following probabilities:

  Letters: {E, L, C, T, R, I, A, N, G}

  Probabilities: {5/21, 2/21, 2/21, 1/21, 2/21, 3/21, 1/21, 3/21, 2/21}

- Code assignment:

							Codeword	Length
E	5/21	0	0			2	00	2
I	3/21	0	1	0		3	010	3
N	3/21	0	1	1		1	011	3
L	2/21	1	0	0		3	100	3
C	2/21	1	0	1	0	4	1010	4
R	2/21	1	0	1	1	2	1011	4
G	2/21	1	1	0		3	110	3
T	1/21	1	1	1	0	4	1110	4
A	1/21	1	1	1	1		1111	4