Motivation

Probability

Reliability: system time to failure is often probabilistic and dependent in analyzable ways on time to failure of components

Algorithm Performance: average case behavior

System Design: if resource demand is probabilistic, what is the distribution of service times?

Simulation: generate probabilistic inputs for a model

Statistics

Analyze data to estimate distributions of probabilistic phenomena, ex. to parametrize models, to assess performance of a system whose behavior differs from trial to trial.

Randomness is difficult if you think too hard about it.

flip a coin ?

digits of π ?

Frequentist: $\lim_{n \to \infty} P\left(\left| \frac{n}{N} - d \right| > \varepsilon \right) = 0$

Probability Theory offers a framework for modeling phenomena with some regularity, but not deterministic predictability

Terminology for modeling nondeterministic situations

A nondeterministic process with an identifiable set containing the possible results is called an experiment.

A single run of an experiment is called a trial.

A set containing all possible results of a random experiment is a sample space for that experiment.

Any single result is an outcome.

An event is a (reasonable) subset of the sample space

Ex. Identify sample spaces for the following experiments. Generally, the sample space for a probability model should be chosen to capture the behavior of interest and to be mathematically tractable.

1) Flip a coin.

2) Roll a die.

3) Run a system until failure.

4) Count hits on a webpage over a 24 hour period.

5) Observe a computer system in operation under a synthetic workload

Sample spaces may be finite, discrete (finite or countably infinite) or continuous, like $\mathbb{R}$, $\mathbb{R} \times \mathbb{R}$, or nice subsets of $\mathbb{R}^n$. There are other possibilities, but they are rare in applications.

Classify the sample spaces chosen for

coin toss
die
failure
hits
system

Any subset of a discrete sample space is a valid event.
Only certain (measurable) subsets of continuous sample spaces are valid events. In practice, it is very rare for a subset of a continuous sample space to be of practical interest but not to be a valid event.

Ex. $\emptyset$, the empty set, is an event for all sample spaces, the impossible event.

If S is a sample space, S is an event.

Likewise, for any $s \in S$, $\{s\}$ is an event.

Def. Events A and B from a sample space S are <u>mutually exclusive</u> if $A \cap B = \emptyset$

Def. A set of events $\{A_\gamma \mid \gamma \in \Gamma\}$ is <u>collectively exhaustive</u> if the union of the events in the collection,

$$\bigcup_{\gamma \in \Gamma} A_\gamma = \{ s \in S \mid \exists \gamma \in \Gamma \, (s \in A_\gamma) \},$$

is S, the sample space.

Def. A set of events $\{A_\gamma \mid \gamma \in \Gamma\}$ is a <u>partition</u> of the sample space S if the set is collectively exhaustive and pairwise mutually exclusive, ie

$$\forall \gamma, \delta \in \Gamma \, (\gamma \neq \delta \rightarrow A_\gamma \cap A_\delta) = \emptyset.$$

Can you think of examples?

Def. A collection of events $\mathcal{E} \subseteq \mathcal{P}(S)$, the set of all subsets of S, that satisfies

1. $S, \emptyset \in \mathcal{E}$
2. $A, B \in \mathcal{E} \Rightarrow A \cup B \in \mathcal{E}$ and $A \cap B \in \mathcal{E}$
3. $A_1, A_2, A_n, \ldots \in \mathcal{E} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{E}$
4. $A \in \mathcal{E} \rightarrow \bar{A} \in \mathcal{E}$

is called a σ-field of S

(Will use $\cap$, $\cup$, and complementation. Remember these? Some properties are listed in §1.5)

Def. A probability space consists of a triple $(S, \mathcal{F}, P)$ where S is a sample space, $\mathcal{F}$ is a σ-algebra of S, and $P: \mathcal{F} \to \mathbb{R}$ is a real valued function on $\mathcal{F}$ satisfying

A1. $\forall A \in \mathcal{F} \quad P(A) \geq 0$

A2. $P(S) = 1$

A3. If A and B are mutually exclusive events then $P(A \cup B) = P(A) + P(B)$

A3'. If $A_1, A_2, \ldots A_n \ldots$ are pairwise mutually exclusive,

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

Ex. $S = \{1, 2, 3, \ldots 13\}$, $\mathcal{F} = \mathcal{P}(S)$

$$P(A) = \frac{|A|}{13}$$ ← cardinality of A

is an example of a probability space with equally likely outcomes. You could use it to model the face value of a card chosen from a well-shuffled standard deck.

Some basic consequences of this definition of a probability space follow:

- If $A_1, A_2, \ldots A_k$ are pairwise mutually exclusive

$$P\left(\bigcup_{i=1}^{k} A_i\right) = \sum_{i=1}^{k} P(A_i)$$

by induction on k:

base case $k=1$

$$P(A_1) = P(A_1) \checkmark$$

Given $P\left(\bigcup_{i=1}^{j} A_i\right) = \sum_{i=1}^{j} A_i$, conclude that

if $A_1, A_2, \ldots A_{j+1}$ are mutually exclusive

because

$B = \bigcup_{i=1}^{j} A_i$ and A_{j+1} are mutually exclusive. $P\left(\bigcup_{i=1}^{j+1} A_i\right) = P(B \cup A_{j+1}) =$

$P(B) + P(A_{j+1})$ by A3; But

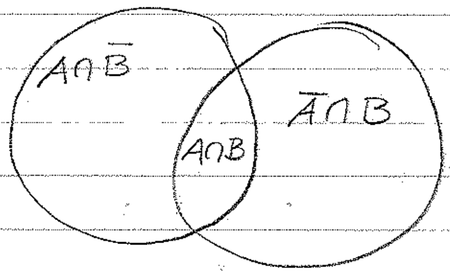
$$P(B) + P(A_{j+1}) = \sum_{i=1}^{j} P(A_i) + P(A_{j+1})$$

$= \sum_{i=1}^{j+1} A_i$, using what we were given about $P\left(\bigcup_{i=1}^{j} A_i\right)$

- $P(\bar{A}) = 1 - P(A)$: $A \cap \bar{A} = \emptyset$ $A \cup \bar{A} = S$ so
  $P(A) + P(\bar{A}) = P(A \cup \bar{A}) = P(S) = 1$

- $P(\emptyset) = 0$, $\emptyset = \bar{S}$

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



$$
\begin{aligned}
P(A \cup B) &= P(A \cap \bar{B}) + P(A \cap B) + P(\bar{A} \cap B) \\
&= \big(P(A \cap B) + P(A \cap B)\big) + \big(P(\bar{A} \cap B) + P(A \cap B)\big) \\
&\quad - P(A \cap B) \\
&= P(A) + P(B) - P(A \cap B)
\end{aligned}
$$

- $A_1, A_2, \ldots A_n \subset S \Rightarrow P\left(\bigcup_{i=1}^{n} A_i\right) =$

$$
\sum_{i=1}^{n} P(A_i) - \sum_{1 \le i < j \le n} P(A_i \cap A_j) + \sum_{1 \le i < j < k \le n} P(A_i \cap A_j \cap A_k) + \ldots
$$

$$
(-1)^{n-1} P\left(\bigcap_{i=1}^{n} A_i\right)
$$

(pf. in text. You can live without it. A Venn diagram for A_1, A_2, A_3 for a finite sample space provides intuition.)

The game, when using probability theory to model a situation, is to construct a probability space that describes the situation, then study the probability space. Iterate through the following steps as necessary:

Identify S.
Assign probabilities, ie identify P.
Identify the subsets of interest, ie events.
Calculate the desired probabilities.

Example: What is the probability that 5 cards drawn from a well shuffled deck will have 4 cards with the same face value?

Sample space S: all possible sets of 5 cards? all possible sequences of 5 cards? (S_1, S_2)

P: For both of the sample spaces under consideration, the outcomes are equally likely, so $s \in S_i$

$$P(\{s\}) = \frac{1}{|S_i|}$$

Subsets of interest
sets of 5 with 4 cards with the same face value : $A_1 \subseteq S_1$

sequences of cards with 4 cards with the same face value : $A_2 \subseteq S_2$

Calculations:

$$|S_2| = 52 \cdot 51 \cdot 50 \cdot 49 \cdot 48 = 311875200$$

$$|S_1| = \frac{52 \cdot 51 \cdot 50 \cdot 49 \cdot 48}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 2598960$$

How many outcomes are in A_1? There are 13 possible face values for the 4 of a kind, and, given the choice of face value, 48 possibilities for the remaining card, so $|A_1| = 13 * 48 = 624$

$$P(A_1) = \frac{624}{2598960} \approx .00024$$

From this, $|A_2| = (5*4*3*2*1)*13*48$

$$\frac{|A_2|}{|S_2|} \approx .00024 \text{, too.}$$

Problem break : p. 19 #2

1.8

In calculating probabilities for a finite sample space with equally likely outcomes, a couple combinatorial formulas are useful.

Ordered Samples of Size k, with Replacement, from a set of n objects: n^k

Ordered Samples of Size k without Replacement from a Set of n Objects

$$n(n-1)\cdots(n-k+1) = \frac{n!}{(n-k)!}$$

Unordered Samples of Size k without Replacement

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

"n choose k".

Ex. Suppose a shipment of 50 capacitors has 5 with actual capacitance outside the required range. If you draw a sample of 10 in which any set of 10 is equally likely to be chosen, what is the probability that all 10 are acceptable

$$|S| = \binom{50}{10} \qquad |A| = \binom{45}{10}, \quad \frac{|A|}{|S|} = \frac{45!}{10!\,35!} \cdot \frac{10!\,40!}{50!}$$

$$= \frac{40\cdot 39\cdot 38\cdot 37\cdot 36}{50\cdot 49\cdot 48\cdot 47\cdot 46} \approx .31$$

Problem break

p. 23 # 4

1.9 Conditional Probability

You may have partial information about the outcome of a trial, the outcome is in event B, say, that affects the probability that the outcome is in event A. The probability of A calculated on the basis of knowledge that the outcome is in B is the probability of A given B, $P(A|B)$.

Def. If $P(B) \neq 0$, $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$

Consequence

$$P(A \cap B) = \begin{cases} P(B)P(A|B) & P(B) \neq 0 \\ P(A)P(B|A) & P(A) \neq 0 \\ 0 & P(A) = P(B) = 0 \end{cases}$$

Example: Suppose that a system can fail with warning or without warning. There are 3 processes, one of which is responsible for a particular failure. Consider the following discrete probability space:

$S = \{(Y,1),(Y,2),(Y,3),(N,1),(N,2),(N,3)\}$

$P(\{(Y,1)\}) = .1$, $P(\{(Y,2)\}) = .1$, $P(\{(Y,3)\}) = .3$
$P(\{(N,1)\}) = .2$, $P(\{(N,2)\}) = .1$, $P(\{(N,3)\}) = .2$

If this is a good model for the situation, and $A = \{(Y,1),(N,1)\}$, outcomes in which component 1 is responsible, while $B = \{(Y,1),(Y,2),(Y,3)\}$, outcomes in which a warning is issued, what is the probability that a warning is issued?
$P(B) = .1 + .1 + .3 = .5$

What is the probability that component 1 is responsible for a particular failure?

$P(A) =$

What is the probability that component 1 is responsible, given that a warning of the failure was issued?

$P(A|B) =$

Problem break:
- If a warning was issued, which component has the highest probability of being responsible? What is that probability?
- What is the probability that component 2 is responsible?
- What is the probability that component 2 is responsible given that a warning was issued?

1.10 Independence of Events

Def. Events A and B are <u>independent</u> if

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

Consequence: If A, B independent and $P(B) \neq 0$ then

$$P(A) = P(A|B) \quad \left(\frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A) \right).$$

The definition of independence captures the intuition that some events are unrelated to one another. Can you think of a real-life situation with two independent events?

Consequence: If A and B are independent, then $\bar{A}$ and B are independent.

$$P(B) = P(A \cap B) + P(\bar{A} \cap B)$$
$$P(B) = P(A)P(B) + P(\bar{A} \cap B)$$
$$P(B) - P(A)P(B) = P(\bar{A} \cap B)$$
$$P(B)(1 - P(A)) = P(\bar{A} \cap B)$$
$$P(B)P(\bar{A}) = P(\bar{A} \cap B) \quad \checkmark$$

Note A, B independent and B, C independent does not imply A, C independent. (Consider $A = C$).

Definition: $A_1, A_2, \ldots A_n$ are independent if for any k, $2 \le k \le n$, and any subset of size k, $\{i_1, i_2, \ldots i_k\} \subseteq \{1, 2, 3 \ldots n\}$,

$$P\left(\bigcap_{j=1}^{k} A_{i_j}\right) = \prod_{j=1}^{k} P(A_{i_j})$$

You really do need to specify all the equalities. Trivedi gives an example due to Ash of events A, B, C satisfying $P(A \cap B \cap C) = P(A)P(B)P(C)$, but not $P(A \cap B) = P(A)P(B)$. There is also an example in which A, B and C are pairwise independent, but $P(A \cap B \cap C) \neq P(A)P(B)P(C)$.

Independence of events is used in basic reliability analysis.

Consider a system with series dependence on its components, For a series system of n components numbered $1, 2, \ldots n$, all the components must be functioning for the system to function. For $i = 1, \ldots n$, let A_i be the event that component i is functioning. Set

$$\text{reliability of } i = R_i = P(A_i)$$

Assuming that the events A_i $i = 1, \ldots n$ are mutually independent, the probability that the system is functioning, R_s, satisfies

$$R_s = \prod_{i=1}^{n} R_i \left(= \prod_{i=1}^{n} P(A_i) = P\left(\bigcap_{i=1}^{n} A_i\right) \text{ by independence.} \right)$$

What is the reliability of a system with 10 components in series with mutually independent reliabilities of .98? .99?

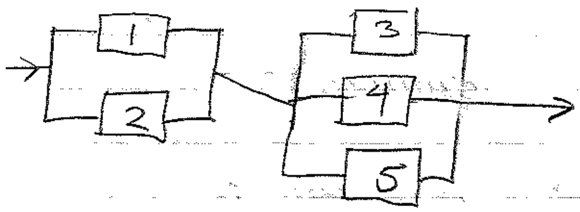
A system has parallel dependence on components numbered 1, 2, ... n if the system will function if at least one of the components is functioning. Compute system reliability, R_s, most conveniently by computing the probability that the system fails and subtracting this from 1.

$$P(\text{system fails}) = \prod_{i=1}^{n} (1 - R_i)$$

$$R_s = 1 - \prod_{i=1}^{n} (1 - R_i).$$

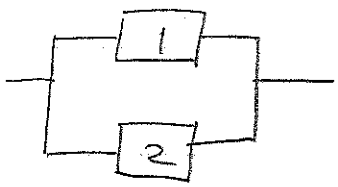
These techniques can be combined hierarchically

Ex.

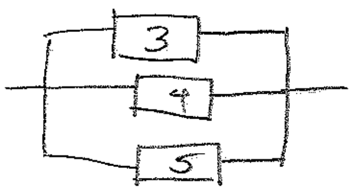


$R_1 = .9 \quad R_2 = .8 \quad R_3 = .95 \quad R_4 = .9 \quad R_5 = .85$

Find R_s.



$= .98$



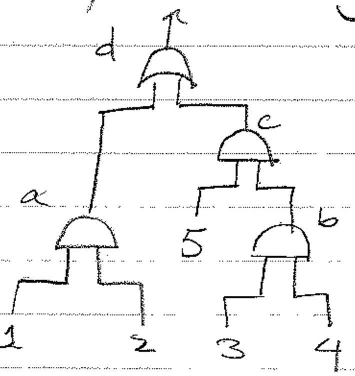
$= .99925$

so $R_s = \quad \approx .979$

System dependence on components may also be represented with a fault tree, a circuit of and ($\cap$) and or ($\Cap$) gates in which the inputs are the failures of the primitive components (T for failure) and the output is the failure of the system.

(Thus $\cap$ corresponds to parallel dependence and $\Cap$ to series.)

For example, the previous system could be represented by



$F_1 = 1-.9 = .1, \quad F_2 = (1-.8) = .2 \quad F_3 = .05, \; F_4 = .1 \quad F_5 = .15$

$F_a = (.1)(.2) = .02, \quad F_b = (.05)(.1), \; F_c = F_b \cdot .15 = .00075$

$F_d = 1-(1-F_a)(1-F_c) \approx 1-.979 \approx .021$

The structure function of a system is a function Φ on bit strings of length n = number of components, say $\vec{x} = (x_1, x_2, \dots x_n)$. $x_i = 0$ if component i is not functioning. $x_i = 1$ if component i is functioning.

$$\Phi(\vec{x}) = \begin{cases} 1 & \text{if system functions under the conditions represented in } \vec{x} \\ 0 & \text{if the system does not function under these conditions} \end{cases}$$

Φ may be given by a truth table or a logical operator.

What is Φ for our pet system?

Ex.

$$\Phi(x_1, x_2, x_3, x_4) = x_1 \wedge (x_2 \vee x_4)) \vee (x_3 \wedge x_4)$$

If x_i is 1 with probability R_i, what is the probability that $\Phi(\vec{x}) = 1$?

Method 1: conditioning on x_1

$$P(\Phi(\vec{x})=1) = P(\Phi(\vec{x})=1 \mid x_1=1)P(x_1)=1 + P(\Phi(\vec{x})=1 \mid x_1=0)P(x_1)=0$$

$$P(\Phi(\vec{x})=1 \mid x_1=1) = P(x_2=1 \vee x_4=1) = 1-(1-R_2)(1-R_4)$$

so

$$P(\underline{\Phi}(\vec{x})=1) = (1-(1-R_2)(1-R_4))R_1 + (R_3R_4)(1-R_1)$$

Method 2: Calculate the probability of each $\vec{x}$ for which $P(\vec{x})=1$ and add

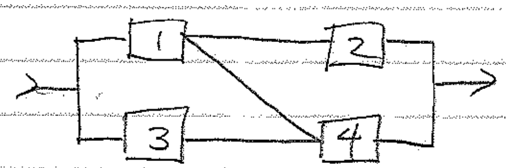
$(x_1, x_2, x_3, x_4) =$

0	0	1	1	$(1-R_1)(1-R_2)R_3R_4$
0	1	1	1	$(1-R_1)R_2R_3R_4$
1	0	0	1	$R_1(1-R_2)(1-R_3)R_4$
1	0	1	1	etc.
1	1	0	0	
1	1	0	1	
1	1	1	1	

$\Big] (1-R_1)R_3R_4$

This can be tedious, but it's mindless.

(Note this system is not series-parallel



)

Problem break p. 35 #2

1.11 Baye's Rule

Given a probability space $(S, \mathcal{F}, P)$,

if $B_1, B_2, \ldots B_n$ is a partition of S and $B_i \in \mathcal{F}$, $1 \le i \le n$, then

$$P(A) = \sum_{i=1}^{n} P(A|B_i)\,P(B_i) \quad \text{and}$$

Bayes' Rule –

$$P(B_j|A) = \frac{P(A|B_j)\,P(B_j)}{\sum_{i=1}^{n} P(A|B_i)\,P(B_i)}$$

$$\left(= \frac{P(A \cap B_j)}{P(A)} \right)$$

The payoff of Bayes' rule is that if we know the $P(B_i)$'s and the $P(A|B_i)$'s, we can turn around and get $P(A)$ and the $P(B_i|A)$'s

Example: A given lot of VLSI chips contains 1% defective chips. Each chip is tested before delivery but the tester is not totally reliable.

$$P(\text{"chip tests as good"} \mid \text{"chip actually is good"})$$

$$= .98$$

$$P(\text{"chip tests as defective"} \mid \text{"chip actually is defective"})$$

$$= .95$$

If a chip selected at random (each chip equally likely to be selected) tests as defective, what is the probability that it actually is defective?

Define events T, (chip tests as defective), and D, (chip is defective).

We want $P(D/T)$

$$= \frac{P(T/D)P(D)}{P(T/D)P(D) + P(T/\bar{D})P(\bar{D})}$$

$$= \frac{(.95)(.01)}{(.95)(.01) + (.02)(.99)} \approx .32$$

Problem break: # 4 p. 45

1.12 Bernoulli trials

A random experiment with 2 outcomes, by convention called "success" and "failure", is called a Bernoulli trial.

$P(\{s\}) = p \qquad P(\{f\}) = q \qquad p+q=1$

If you perform n independent repetitions of the same Bernoulli process, the probability of k successes is

$\binom{n}{k} p^k q^{n-k}$.

Ex. TMR, triple modular redundancy: A system has 3 components, 2 of which must be functioning for the system to function. Let the reliability of the components be R.

$$R_{TMR} = \sum_{i=2}^{3} \binom{3}{i} R^i (1-R)^{3-i}$$

$$= 3R^2(1-R) + R^3 = 3R^2 - 2R^3$$

Note $R_{TMR} - R = -2R^3 + 3R^2 - R = R(-2R^2 + 3R - 1)$ with zeros at 0, ½, and 1. Check that

$$R_{TMR} \begin{cases} > R & R > \frac{1}{2} \\ = R & R = \frac{1}{2} \\ < R & R < \frac{1}{2} \end{cases}$$

Generalized Bernoulli Trials:

A generalized Bernoulli trial has k possible outcomes, $b_1, b_2, \ldots b_k$ with probabilities $p_1, p_2, \ldots p_k$ satisfying

$$\sum_{i=1}^{k} p_i = 1.$$

Given n independent identically distributed generalized Bernoulli trials and k indices $n_1, n_2, \ldots n_k$, $0 \le n_i \le n$, $\sum_{i=1}^{n} n_i = n$,

the probability that each b_i will occur exactly n_i times is

$$\frac{n!}{\prod_{i=1}^{k} n_i!} \prod_{i=1}^{k} p_i^{n_i} = \binom{n}{n_1\, n_2 \ldots n_k} \prod_{i=1}^{k} p_i^{n_i}$$

(Intuition: $k=3$. There are $\binom{n}{n_1}$ positions for n_1 b_1 in a sequence of n trials. There remain $n-n_1$ locations for n_2 b_2's in each of these arrangements of b_1's. Thus there are

$$\binom{n}{n_1}\binom{n-n_1}{n_2} = \frac{n!}{n_1!\,\cancel{(n-n_1)!}} \frac{\cancel{(n-n_1)!}}{n_2!\,(n-n_1-n_2)!}$$

$$= \frac{n!}{n_1!\,n_2!\,n_3!}$$ arrangements of n_1 b_1's

n_2 b_2's and n_3 b_3's