MAT2377

Rafał Kulik

Version 2015/September/15

Comments

- These notes cover material from Chapter 1, Sections 1.1-1.9. **In class, I will use a blackboard**.
- I'm not re-writing the textbook.
- In class I do some examples that are not included in the notes and vice versa: here you can find some examples that I do not do in class.
- There may be some typos. The final version of the slides will be posted *after* the chapter is finished.

Sample spaces and events

We will deal with **random experiments** (e.g. measurements of speed/weight, number and duration of phone calls).

For any "experiment" define the **sample space** as the set of all possible outcomes. This is often denoted by the symbol $\mathcal{S}$.

A sample space can be discrete or continuous.

An **event** is a collection of outcomes from the sample space $\mathcal{S}$. Events will be denoted by A, B, E_1, E_2 etc.

Examples:

- Toss a fair coin. The (discrete) sample space is $\mathcal{S} = \{\text{Head, Tail}\}$.

- Roll a die: The (discrete) sample space is $\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$. Various events:
  - Roll an even number: represent this as $\{2, 4, 6\}$.
  - Roll a prime number: $\{2, 3, 5\}$.

- Suppose we measure the weight (in grams) of a chemical sample. The (continuous) sample space can be represented by $(0, \infty)$, the positive half line. Events:
  - sample is less than 1.5 grams: $(0, 1.5)$;
  - sample exceeds 5 grams: $(5, \infty)$;

For any events A and B in $\mathcal{S}$:

- The Union of A and B are all outcomes from $\mathcal{S}$ in either A or B; write as $A \cup B$.
- The Intersection of A and B are all outcomes in both A and B; write $A \cap B$.
- The Complement of A is the set of all outcomes in $\mathcal{S}$ that are not in A. Write A^c for this.
- If A and B have no outcomes in common, they are mutually exclusive; write $A \cap B = \emptyset$ (the empty set). In particular, A and A^c are mutually exclusive.
- Graphical representation of events - Venn diagrams $\Rightarrow$ blackboard.

Examples:

- Roll a die. Let $A = \{2,3,5\}$ (a prime number) and $B = \{3,6\}$ (multiple of 3). Then $A \cup B = \{2,3,5,6\}$, $A \cap B = \{3\}$ and $A^c = \{1,4,6\}$.

- 100 samples of plastic are analyzed for scratch and shock resistance.

		shock resistance	
		high	low
scratch resistance	high	40	4
	low	1	55

If A is the event that a sample has high shock resistance and B is the event that a sample has high scratch residence, then $A \cap B$ consists of 40 samples.

Counting techniques

Note: This slide is not discussed in class

We now consider some basic combinatorial results which make later some probabilities easier to calculate.

A **two-stage Procedure** procedure can be modeled as having k bags, with m_1 items in the first bag,. . . , m_k items in k-th bag.

The first stage consists of picking a bag, and the second stage consists of drawing an item out of that bag.

But this is equivalent to picking one of the $m_1 + m_2 + \cdots + m_k$ total number of items.

If all the bags have the same number $m_1 = \cdots = m_k = n$ then there are kn items in total, and this is the total number of ways the two-stage procedure can happen.

Examples

Note: This slide is not discussed in class

- How many ways can I roll a die and then draw a card from a shuffled 52-card pack?
  - There are 6 ways the first step can turn out, and for each of these (the stages are independent in fact) there are 52 ways to draw the card. Thus there are $6 \times 52 = 312$ ways this can turn out.
- How many ways can I draw out two tickets numbered 1 to 100 from a bag: first right-hand and then left hand?
  - There are 100 ways to pick the first number; for *each of these* there are 99 ways to pick the second number. Thus 9900 ways.

Any number of stages

Note: This slide is not discussed in class

This leads us to general multi-stage procedures: If we have a k-stage process like this where there are n_1 possibilities for stage 1, regardless of the 1st outcome there are n_2 possibilities at stage 2, etc., up to n_k choices at stage k, then there are

$$n_1 n_2 \cdots n_k$$

total ways the process can turn out.

Ordered samples

Suppose we have items numbered $1, 2, \ldots, n$ and are drawing an ordered sample of size r of them

- **with replacement**;
- **without replacement**.

How many ways can each case turn out?

Each is an r-stage procedure.

With Replacement (order important)

If we replace each "number" after it is drawn, then every draw is the same and there are n ways it can turn out.

So according to our earlier result there are

$$\underbrace{n\,n\cdots n}_{r \text{ factors}} = n^r$$

ways to choose an ordered sample of size r with replacement from $\{1, 2, \ldots, n\}$.

Without Replacement (order important)

If we **do not** replace each "number" after it is drawn, then choices for second draw depend on first draw but regardless there are $(n-1)$ choices.

Also, whatever the first two draws, there are $(n-2)$ ways to draw the third number, etc.

Thus there are

$$\underbrace{n(n-1)\cdots(n-r+1)}_{r \text{ factors}} = {}_nP_r \quad \text{(common calculator symbol)}$$

ways to choose an ordered sample of size $r \leq n$ without replacement from $\{1, 2, \ldots, n\}$.

Factorial notation

Writing $n! = n(n-1)(n-2)\cdots 1$, for positive integer n, we have

- when $r = n$, ${}_nP_r = n!$, and the ordered selection is called a **permutation**;
- when $r < n$, we can write

$$ {}_nP_r = \frac{n(n-1)\cdots(n-r+1)}{}\frac{(n-r)\cdots 1}{(n-r)\cdots 1} = \frac{n!}{(n-r)!}\,. \tag{1} $$

If *by convention* we take $0! = 1$ then equality (1) is also true for $r = n$.

Examples:

1. How many different ways can 6 balls be drawn *in order* without replacement from balls numbered 1 to 44?

Answer: ${}_{44}P_6 = 44 \times 43 \times 42 \times 41 \times 40 \times 39 = 5,082,517,440.$

This is the number of ways the actual drawing of the balls could occur (e.g. watching the draw on TV).

2. Having digits $\{0, 1, \ldots, 9\}$, how many 6-digits PIN codes can you create

- if digits may be repeated: $10\times10\times10\times10\times10\times10 = 10^6 = 1,000,000.$

- if digits may not be repeated: ${}_{10}P_6 = 10 \times 9 \times 8 \times 7 \times 6 \times 5.$

Unordered sample

Consider sampling without replacement. Now suppose that we cannot distinguish between different ordered samples that contain the same elements (e.g. looking up Lotto results in the newspaper). Denote the (as yet unknown) number of unordered samples by ${}_nC_r$.

We can deduce this by noting that the following two processes are equivalent:

- Take an ordered sample of size r.
- Take an unordered sample of size r **and then** rearrange (permute) the numbers.

There are ${}_nP_r$ ways to perform the first procedure.

The second is a two-stage process. There are ${}_nC_r$ ways to do the first stage. For each of these there are $r!$ ways to rearrange/permute the r numbers.

Thus

$$ {}_nP_r = C_r^n \times r! \quad \Rightarrow \quad {}_nC_r = {}_nP_r / r! = \frac{n!}{(n-r)!r!} = \binom{n}{r}. $$

This last notation is common in many texts, and is also called a **binomial coefficient**. Read as "n-choose-r".

Example:

How many ways can the "Lotto draw" be reported in the newspaper (where they are always reported in increasing order)?

This is the same as the number of *unordered samples* (different reorderings of same 6 numbers are indistinguishable), and so

Answer:

$$ {}_{44}C_6 = \binom{44}{6} = \frac{44 \times 43 \times 42 \times 41 \times 40 \times 39}{6 \times 5 \times 4 \times 3 \times 2 \times 1} = \frac{5,082,517,440}{720} = 7,059,052\,. $$

Probability of an Event

For situations where we have a random experiment which has exactly c possible mutually exclusive, equally likely, simple outcomes we can assign a probability to an event A by counting the number of simple outcomes that correspond to A. If the count is a then

$$P(A) = \frac{a}{c}.$$

$\triangleright$ if a sample space consists of N equally likely events, the probability of each outcome is $1/N$.

Example 1:

1. Toss a fair coin. The sample space is $\mathcal{S} = \{\text{Head, Tail}\}$ The probability of observing a Head is $\frac{1}{2}$.

2. Throw a fair six sided die. There are 6 possible outcomes.

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

If A corresponds to observing a multiple of 3 then, in set notation

$$A = \{3, 6\}.$$

$$\text{Prob(number is a multiple of 3)} = P(A) = \frac{2}{6} = \frac{1}{3}.$$

Furthermore:

- Prob$\{$even no.$\} = P(\{2,4,6\}) = \frac{3}{6} = \frac{1}{2}$.

- Prob$\{$prime no.$\} = P(\{2,3,5\}) = 1 - P(\{1,4,6\}) = \frac{1}{2}$.

3. In a group of 1000 people it is known that 545 have high blood pressure. 1 person is selected randomly. What is the probability that this person has high blood pressure?

Answer: Relative frequency of people with high blood pressure is $\frac{545}{1000}$. Via the classical definition, this is the probability we are looking for.

Axioms of probability

1. For any event A, $1 \geq P(A) \geq 0$.

2. For the complete sample space $\mathcal{S}$, $P(\mathcal{S}) = 1$.

3. For two mutually exclusive events A and B the probability that A or B occurs is $P(A \cup B) = P(A) + P(B)$.

Since $\mathcal{S} = A \cup A^c$, and A, A^c mut. excl., $1 \stackrel{\textbf{A2}}{=} P(\mathcal{S}) = P(A \cup A^c) \stackrel{\textbf{A3}}{=} P(A) + P(A^c)$. Thus

$$P(A^c) = 1 - P(A)$$

Example 2: 1. Throw a single six sided die. Let $A = \{3, 6\}$ - the number is a multiple of 3, $B = \{1, 2\}$ - the number is less than 3. A and B are mutually exclusive.

$$P(A \text{ or } B \text{ occurs}) = P(A) + P(B) = \frac{2}{6} + \frac{2}{6} = \frac{2}{3}.$$

2. I have an urn containing 4 white balls, 3 red balls and 1 blue ball. I draw one ball and note the events $W = \{\text{the ball is white}\}$, $R = \{\text{the ball is red}\}$ and $B = \{\text{the ball is blue}\}$. Then

$$P(W) = 4/8 = 1/2, \quad P(R) = 3/8, \quad P(B) = 1/8.$$

Also, the probability of drawing a white or a red is

$$P(W \text{ or } R) = P(W \cup R) = 7/8.$$

General addition rule

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

Example 3: An electronic assembly consists of two components A and B. Experience tells us that $P\{A \text{ fails}\} = 0.2$, $P\{B \text{ fails}\} = 0.3$ and $P\{\text{both } A \text{ and } B \text{ fail}\} = 0.15$.

Find $P\{\text{at least one of } A \text{ and } B \text{ fails}\}$ and $P\{\text{neither } A \text{ nor } B \text{ fails}\}$. Write A for "A fails" and sim'ly for B. Then we want

$$\begin{aligned} P\{\text{at least one fails}\} &= P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.35\,; \\ P\{\text{neither fail}\} &= 1 - P\{\text{at least one fails}\} = 0.65\,. \end{aligned}$$

More than two events:

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

Patient: "Will I survive this risky operation?"
Surgeon: "Yes, I'm absolutely sure that you will survive the operation."
Patient: "How can you be so sure?"
Surgeon: "9 out of 10 patients die in this operation, and yesterday died my ninth patient."

One out of every four people is suffering from some form of mental illness. Check three friends. If they're OK, then it's you.

Conditional probability and independent events

Any two events A and B satisfying

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

are said to be **independent**; this is a purely mathematical definition, but agrees with intuitive notion in simple examples.

- Flip a coin twice, so 4 possible outcomes are equally likely: let $A = \{HH\} \cup \{HT\}$ denote "head on first flip", $B = \{HH\} \cup \{TH\}$ "head on second flip". Then

$$P(A) = P(HH) + P(HT) = \frac{1}{2} \quad , \quad P(B) = P(HH) + P(TH) = \frac{1}{2}.$$

But $P(A \cap B) = P(HH) = \frac{1}{4} = P(A)P(B)$, so A and B are independent.

- A card is drawn from a well shuffled deck. Let A be the event that it is an ace and let D be the event that it is a diamond.

To check: $P(A) = \frac{1}{13}$ and $P(D) = \frac{1}{4}$. Also $P(A \cap D) = \frac{1}{52} = P(A)P(D)$.

- A six-sided die numbered 1–6 is loaded in such a way that the prob of getting each value is *proportional* to that value. Find prob. of a 3.

  For some value v, $P(1) = v$, . . . , $P(6) = 6v$. Since these add to 1,

$$1 \quad = \quad v + 2v + 3v + 4v + 5v + 6v = 21v\,.$$

  Hence $v = 1/21$, $P(3) = 3v = 3/21 = 1/7$. Now the die is rolled twice *independently*. Find the prob of getting two 3's.

  If experiment is such that $P(\text{3 on lst}) = 1/7$ and $P(\text{3 on 2nd}) = 1/7$ **and** the two events are independent, then

$$P\left(\{\text{3 on 1st}\} \cap \{\text{3 on 2nd}\}\right) = P(\text{3 on 1st})P(\text{3 on 2nd}) = 1/49\,.$$

- Which plane is more likely to crash: 2- or 3- engines one?

  Answer: Assumptions - engines fail independently. We assume that the probability that the engine fails is p. A 2 engine plane will crash if both engines fail - probability p^2. A 3 engine plane will crash if any pair of engines fail, or if all 3 fail together. The probability of a pair of engines failing is $p * p * (1 - p)$: i.e. FAIL FAIL OK. There are 3 DIFFERENT pairs to be considered: AB, BC, or AC. The probability of all three engines failing is p^3. Therefore the probability of at least 2 engines failing is: $3p^2(1 - p) + p^3 = 3p^2 - 2p^3$. So, what does all this mean? Well, basically it's safer to use a 2-engine plane than a 3-engine plane: the 3-engine plane will crash more often, assuming that it needs 2 engines to fly.

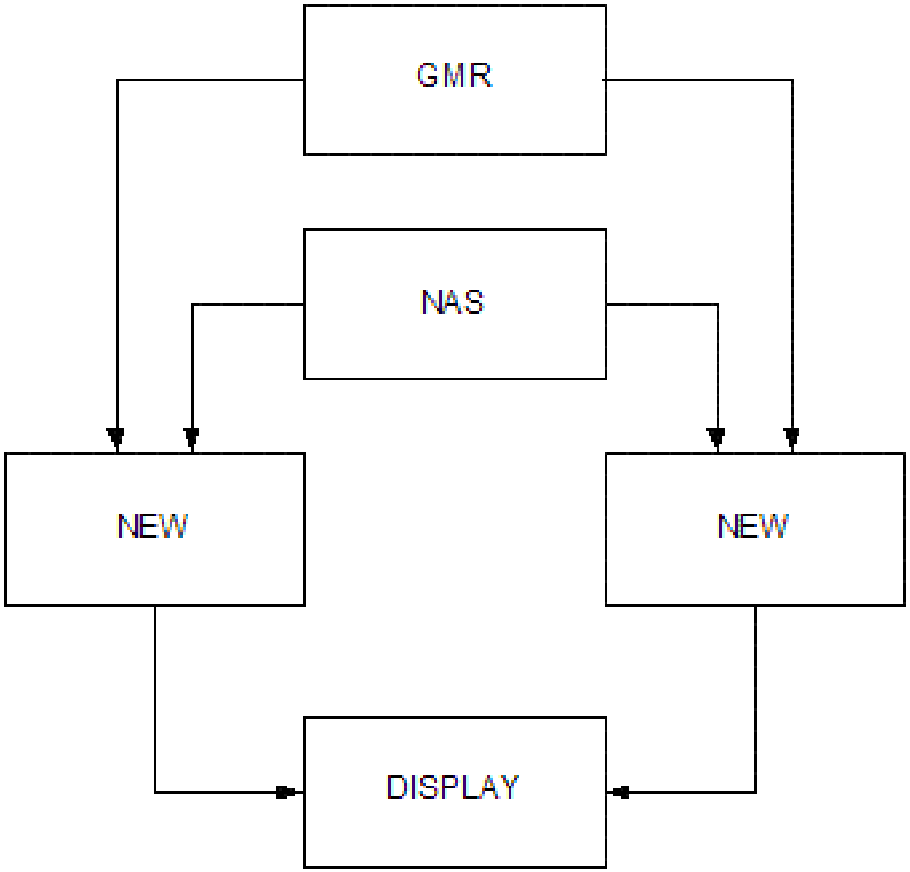
  You can sort of make sense of this by thinking that the 2-engine plane needs 50% of its engines working, while the 3-engine plane needs 66%. Of course, you could always travel by Greyhound.

Air Traffic Control Example

Air Traffic Control is a Safety Related activity. Each piece of equipment is therefore designed to the highest safety standards and in many cases duplicate equipment is provided so that if one item fails another takes over.

A new system is to be provided passing information from Heathrow Airport to the Terminal Control Room at West Drayton. As part of the system design a decision has to be made as to whether it is necessary to provide duplication.

The new system takes data from the *Ground Movements Radar* (GMR) at Heathrow, combines this with data from *NAS*, and sends the output to a display in the *Terminal Control Room*.


GMR
NAS
NEW
NEW
DISPLAY

For all existing systems records of failure are kept and an experimental probability of failure is calculated annually using data from the previous 4 years. The *reliability* of a system is then defined as $R = 1 - P$, where $P = P(failure)$.

We know that $R_{GMR} = R_{NAS} = 0.9999$.

For the system above: If a single New system is introduced the reliability of the system is $R_{GMR} \times R_{NEW} \times R_{NAS}$. If the New system is duplicated the reliability of the system is $R_{GMR} \times (1 - (1 - R_{NEW})^2) \times R_{NAS}$ Duplicating the New system therefore causes an improvement in reliability of

$$(1 - (1 - R_{NEW})^2) \times 100\%/R_{NEW}.$$

For the new system, no historical data is available so we work out the improvement achieved by using the dual thread design for various values of R_{NEW}.

R_{NEW}	0.1	0.2	0.5	0.75	0.99	0.999	0.9999	0.99999
Improvement	190	180	150	125	101	100.1	100.01	100.001

If the new system was very unreliable then there would be a significant benefit in using the dual thread design.

However we have no intention of installing an unreliable system.

If the new system is as reliable as NAS and GMR i.e.$R_{GMR} = R_{NEW} = R_{NAS} = 0.9999$, then the single thread design would give a reliability of 0.9997 (equals 3 failures in 10,000 hours) whereas the dual thread design would give a reliability of 0.9998 (equals 2 failures in 10,000 hours)

We could conclude from this that the reliability gain from a dual thread design would not justify the extra cost.

Conditional Probability

We can better understand independence by defining the conditional probability of $\boldsymbol{B}$ given $\boldsymbol{A}$ as

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

Note that this only makes sense when "A can happen" i.e. $P(A) > 0$.

Example: From a group of 100 people, 1 is selected. What is the probability that this person has high blood pressure? This is (unconditional) probability.

Now, from this group we select first all people with high cholesterol level, and then from the latter group we select 1 person. What is the probability that this person has high blood pressure? This is conditional probability; the probability of selecting a person with high blood pressure, given high cholesterol level.

Example: A sample of 249 Eskimos was taken and each person classified by blood group and tuberculosis(TB) status.

	O	**A**	**B**	**AB**	Total
TB	34	37	31	11	113
No TB	55	50	24	7	136
Total	89	87	55	18	249

The rel. frequency of people with TB is $\frac{113}{249} = 0.454$. In other words, the prob. that randomly selected person has TB is 0.454.

Of those with type **B** blood the rel. frequency of TB is

$$\text{rel.fr(TB}|\text{type } \mathbf{B}) = \frac{31}{55} = \frac{31/249}{55/249} = 0.564.$$

Total Probability Rule

If $A_1, ...A_k$ are mutually exclusive and exhaustive (i.e. $A_i \cap A_j = \emptyset$ if $i \neq j$ and $A_1 \cup .... \cup A_k = S$), then for any event B

$$P(B) = P(B \mid A_1)P(A_1) + ... + P(B \mid A_k)P(A_k)$$

Bayes' Theorem

After an experiment generates an outcome, we are interested in the probability that a condition was present given an outcome.

If $A_1, ...A_k$ are mutually exclusive and exhaustive (i.e. $A_i \cap A_j = \emptyset$ if $i \neq j$ and $A_1 \cup .... \cup A_k = S$), then for any event B and for each i,

$$P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{P(B)} = \frac{P(B \mid A_i)P(A_i)}{P(B \mid A_1)P(A_1) + ... + P(B \mid A_k)P(A_k)}$$

Example: Nissan sold three models of cars in North America in 1999: Sentras, Maximas and Pathfinders. Of the vehicles sold, 50% were Sentras, 30% were Maximas and 20% were Pathfinders. In the same year 12% of the Sentras, 15% of the Maximas and 25% of the Pathfinders had a defect.

1. I own a 1999 Nissan. What is the probability that it has the defect?

2. My 1999 Nissan has the defect. What model do you think I own? Why?
   Answer: In the first part we want to compute the total probability $P(D)$:
   In the second part we compare $P(M|D)$, $P(S|D)$, $P(Pa|D)$:

$$P(M|D) = \frac{P(D|M)P(M)}{P(D)} = ...$$