UNIVERSITÉ Concordia UNIVERSITY

COMP 233/2

Probability and Statistics for Computer Science

Week 2

Conditional probability,
Bayes' formula,
Independent events

Reading: Chap 4

Food for thought – Example 1.1

- A student is taking a Data Structures test in which 7 questions out of 10 must be answered. In how many ways can the student answer the exam if
  1. Any 7 questions may be selected.
  2. The first two questions must be selected.
  3. The student must choose three questions from the first five and four questions from the last five.

"7 questions out of 10" Solution – Part 1).

Any 7 questions may be chosen.

- Thus, the student needs to select 7 objects out of a set of 10 distinct objects.
- Hence, the student can answer the exam in $\binom{10}{7}$ different ways.
- So the answer is 10!/(3!7!) = 120.

Solution - Part 2).

The first two questions must be chosen.

- We know that 2 of the selected 7 are the first two.
- Thus, different selections correspond to different selections of the remaining 5 out of only 8 questions.
- Hence, the student can answer the exam in $\binom{8}{5}$ = 8!/(3!5!) = 84 different ways.

Solution – Part 3).

3 selected from the first 5 and 4 selected from the last 5.

- The selection is actually a sequence of two selections: 3 from 5 and 4 from 5.
- I.e., $\binom{5}{3}$ possible selections followed by $\binom{5}{4}$ possible selections
- Thus, by the rule of products, the student can answer the exam in $10 \times 5 = 50$ different ways.

Food for thought - Example 1.2

- Show that $$\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}$$

$$\binom{n-1}{r} + \binom{n-1}{r-1}$$

(n-1)!/(n-1-r)!r! + (n-1)!/(n-1-r+1)!(r-1)!

= (n-1)!/(n-r-1)!r! + (n-1)!/(n-r)!(r-1)!

- Note that (n-r)! = (n-r)(n- 1 - r)!
- So above equals

(n-r)(n-1)!/(n-r)!r! + r(n-1)!/(n-r)!r!

= [(n-r)(n-1)! + r(n-1)!]/(n-r)!r!

= n!/(n-r)!r!

Q.E.D.

- This is the basis of Pascal's triangle.

Food for thought - Example 1.3

- Show that

$$2^n = \sum_{r=0}^{n} \binom{n}{r}$$

- Can you give a probabilistic reasoning that proves the above property?

- "n choose r" is the number of different subsets consisting of r distinct objects from n.
- Let us consider n = 3

0	0	0	for r = 0
0	0	1	for r = 1
0	1	0	
1	0	0	
0	1	1	for r = 2
1	0	1	
1	1	0	
1	1	1	for r = 3

- Each object is chosen (1) or not (0). So, 2 choices
- By multiplication rule it is $2*2*2 = 2^3$

Food for thought – Example 1.4

- Ten students (A, B, . . ., J) are to be divided into five teams of two to work on a term project.
  a) Find the number of possible divisions into teams.
  b) Find the probability that A and B form a team.

Solution to "10 students divided into 5 teams"

a) Here the sample space is the set of all possible pairings into teams.
   - Each element in the sample space is a set of five pairs.
   - $N(S) = (10)!/((2!)(2!)(2!)(2!)(2!)(5!))$
     $= (10)!/((2^5)(5!))$

b) E is the set of pairings in which A and B form a pair.
   - $N(E) = (8!)/((2^4)(4!))$
   - Hence, get $P(E) = N(E)/N(S)$.

Example

Two cards are drawn from a standard deck and lined up on a table. Find the probability that the first (leftmost) card is a king, and the second one is not.

Solution:

- Here S is the set of all permutations of 2 cards. Thus, $N(S) = {}^{52}P_2 = 52 \times 51$.
- The number of outcomes in E is – by the multiplication rule – $N(E) = 4 \times 48$ (4 kings for the first card and 48 not-kings for the second one).
- Hence, $P(E) = N(E)/N(S) = (4 \times 48)/(52 \times 51)$.

A closer look

$$P(E) = \frac{4 \cdot 48}{52 \cdot 51} = \frac{4}{52} \cdot \frac{48}{51}$$

$$E = KN$$

$$P(K) = \frac{4}{52}, \; so$$

$$P(KN) = P(K) \cdot P(???)$$

One event affects the probability of another event

- The mysterious probability is the so-called **conditional probability** that the second card (N_2) is not a king given that the first one (K_1) is a king, i.e.

$$P(K_1N_2) = P(K_1) \cdot P(N_2 \mid K_1),$$

$$or,$$

$$P(N_2 \mid K_1) = \frac{P(K_1N_2)}{P(K_1)}$$

Example

- In a factory, 400 parts are taken for inspection on whether they have surface flaws, and whether they are [functionally] defective.

Df \ SF	Yes	No	Total
Yes	10	18	28
No	30	342	372
Total	40	360	400

Example

What is the probability that

1. **a part is defective, if it has a surface flaw?**
2. a part is defective, if it has no surface flaw?

Df \ SF	Yes	No	Total
Yes	10	18	28
No	30	342	372
Total	40	360	400

Solution

- Let D and F be the events that a part is defective and that it has a surface flaw, respectively
- We observe that $N(S) = 400$, $N(D) = 28$, $N(F)=40$.
- Hence, $P(D) = 28/400 = 0.07$ and
  $P(F) = 40/400 = 0.1$

Df \ SF	Yes	No	Total
Yes	10	18	**28**
No	30	342	372
Total	**40**	360	**400**

Solution

- If a part has a surface flaw, then F occurred. In other words, we only look at 40 parts with surface flaws.
- We observe that $N(D \text{ and } F) = N(DF) = 10$, $N(F) = 40$. Hence, $P(D \mid F) = 0.25$.
- *Note: the above probability is not the probability of "D and F"! That one equals 10/400 = 0.025.*

Df \ SF	Yes	No	Total
Yes	**10**		
No	30		
Total	**40**		

Reminder

What is the probability that

1. a part is defective, if it has a surface flaw?
2. **a part is defective if it has no surface flaw?**

Df \ SF	Yes	No	Total
Yes	10	18	28
No	30	342	372
Total	40	360	400

Solution

- If a part has no surface flaw, then F^c occurred. In other words, we only look at 360 parts without surface flaws.
- We observe that $N(D \text{ and } F^c) = N(DF^c) = 18$, $N(F^c) = 360$. Hence, $P(D \mid F^c) = 0.05$.
- *Note: the above probability is not the probability of the intersection "D and F^c"! That one equals 18/400 = 0.045.*

Df \ SF	Yes	No	Total
Yes		**18**	
No		342	
Total		**360**	

Summary of Example

$$P(D)=\frac{N(D)}{N(S)}=\frac{28}{400}=0.07$$

$$P(F)=\frac{N(F)}{N(S)}=\frac{40}{400}=0.1$$

$$P(D\mid F)=\frac{N(DF)}{N(F)}=\frac{10}{40}=0.25$$

$$P(D\mid F^{c})=\frac{N(DF^{c})}{N(F^{c})}=\frac{18}{360}=0.05$$

Example (revisited)

What is the probability that

1. a part is defective, if it has a surface flaw?
2. a part is defective if it has no surface flaw?

	F	F^c	Prob
D	0.025	0.045	
D^c	0.075	0.855	
Prob			

$P(DF^c)$

Example (revisited)

What is the probability that

1. a part is defective, if it has a surface flaw?
2. a part is defective if it has no surface flaw?

	F	F^c	Prob
D	0.025	0.045	0.07
D^c	0.075	0.855	0.93
Prob	0.1	0.9	

Summary of Example

$$P(D) = P(DF) + P(DF^c) = 0.07$$

$$P(F) = P(FD) + P(FD^c) = 0.1$$

$$P(D \mid F) = \frac{P(DF)}{P(F)} = \frac{0.025}{0.1} = 0.25$$

$$P(D \mid F^c) = \frac{P(DF^c)}{P(F^c)} = \frac{0.045}{0.9} = 0.05$$

Conditional Probability

- Let F be an event such that $P(F)>0$. The ***conditional probability of an event E, given that F occurred*** is given by

$$P(E \mid F) = \frac{P(EF)}{P(F)}$$

Example

- A bin contains 5 defective (that immediately fail when put in use), 10 partially defective (that fail after a couple of hours of use), and 25 acceptable transistors.
- A transistor is chosen at random from the bin and put into use. If it does not immediately fail, what is the probability it is acceptable?

Example - Analysis

- A bin contains 5 defective (that immediately fail when put in use), 10 partially defective (that fail after a couple of hours of use), and 25 acceptable transistors.

	D	D^c	Total
A			
A^c			
Total			

Solution

- Since the transistor did not immediately fail, we know that it is not one of the 5 defectives and so the desired probability is:

$$P\{\text{acceptable} | \text{not defective}\} = \frac{P\{\text{acceptable, not defective}\}}{P\{\text{not defective}\}} = \frac{P\{\text{acceptable}\}}{P\{\text{not defective}\}}$$

where the last equality follows since the transistor will be both acceptable and not defective if it is acceptable.

Solution

- Hence, assuming that each of the 40 transistors is equally likely to be chosen, we obtain that

P{acceptable|not defective} = [25/40]/[35/40] = 5/7

- It should be noted that we could <u>also</u> have derived this probability by working directly with the reduced sample space. That is, since we know that the chosen transistor is not defective, the problem reduces to computing the probability that a transistor, chosen at random from a bin containing 25 acceptable and 10 partially defective transistors, is acceptable. This is clearly equal to 25/35 .

Intersection via Conditional Probability

- Sometimes – think of an example with drawing two cards without replacement – it is easy to find conditional probabilities, whereas computing the probability of the intersection is more challenging. Then

$$P(EF) = P(E \mid F)P(F)$$

Example

- Guy Lemieux figures that there is a 25 percent chance that he will be traded from his current team to the Montréal Canadiens. If he is traded, he is 50 percent certain that he will be made Captain of the Canadiens. What is the probability that Guy will be Captain of the Montréal Canadiens?

Solution

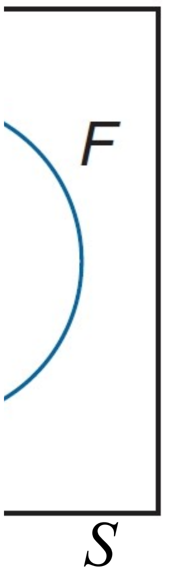
- If we let T denote the event that Guy is traded to the Canadiens, and C the event that Guy is made Captain of the Canadiens, then the desired probability is P(TC), which is obtained as follows:

$$P(TC) = P(T)P(C|T) = (.25)(.5) = .125$$

- Hence, there is an 12.5 percent chance that Guy will be Captain of the Canadiens.

Observation

- For every event F, the sample space can be represented as $S = F \cup F^c$. Then for every event E we have



$$
\begin{aligned}
P(E) &= P(ES) \\
&= P(E(F \cup F^c)) \\
&= P((EF) \cup (EF^c)) \\
&= P(EF) + P(EF^c) \\
&= P(E|F)\,P(F) + P(E|F^c)\,P(F^c)
\end{aligned}
$$

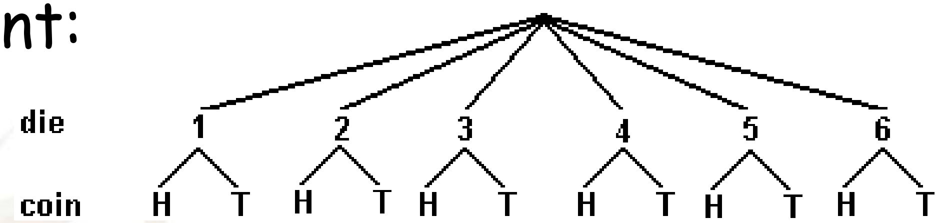
Total Probability Rule

- We have just discovered the so-called total probability rule.
- It allows one to express the probability of any event via certain conditional probabilities.

$$P(E) = P(E \mid F)P(F) + P(E \mid F^c)P(F^c)$$

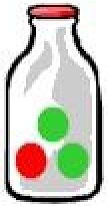
Probability Tree Diagrams

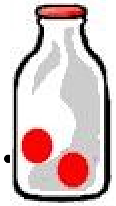
- A **probability tree diagram** shows all the possible events of an experiment. The first event is represented by a dot. From the dot, branches are drawn to represent all possible outcomes of that event (typically each branch shows its probability).
- Example: Consider an experiment where a die is cast followed by flipping a coin. The tree diagram charts of all the possible outcomes of the experiment:



http://cnx.org/content/m19463/latest/

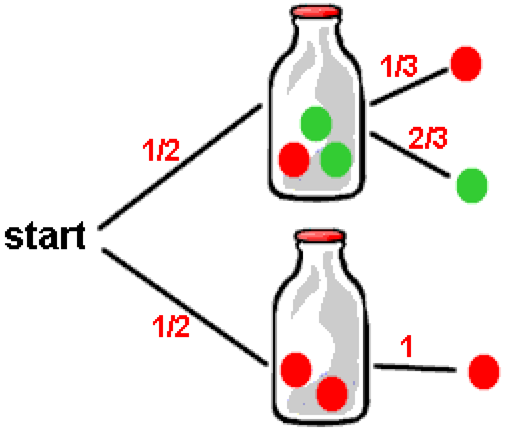
Example

- There are two identical bottles.
- One contains 2 green balls and 1 red ball, 

  the other bottle contains 2 red balls. 
- Experiment: A bottle is selected at random and then a ball is drawn from it.
- What is the probability that the ball is red?

Solution (with tree diagram)

- Let I and II stand, respectively, for the events that the first bottle and the second bottle were selected. Hence, $P(I) = P(II) = 0.5$. The chances to draw a red ball from Bottle 1 are $1/3$. For Bottle 2 these chances are 1. So…

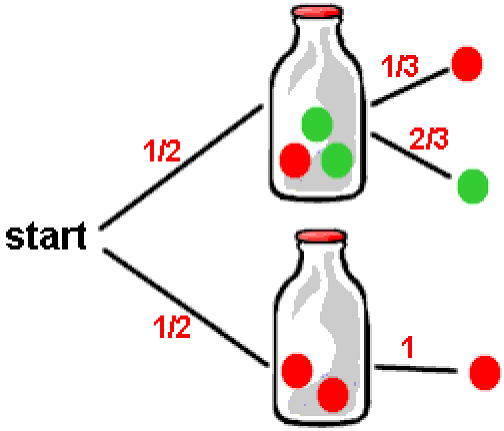


(tree diagram)

Solution

- by the *total probability rule* the requested probability is

$$
\begin{aligned}
&P(\text{Red}) \\
&= P(\text{Red} \mid \text{I})\,P(\text{I}) + P(\text{Red} \mid \text{II})\,P(\text{II}) \\
&= \frac{1}{3}\cdot\frac{1}{2} + 1\cdot\frac{1}{2} = \frac{1}{6} + \frac{1}{2} = \frac{2}{3}
\end{aligned}
$$



Total Probability Rule: General Case.

- Suppose that S is represented as the union of n disjoint events F_k. That is,

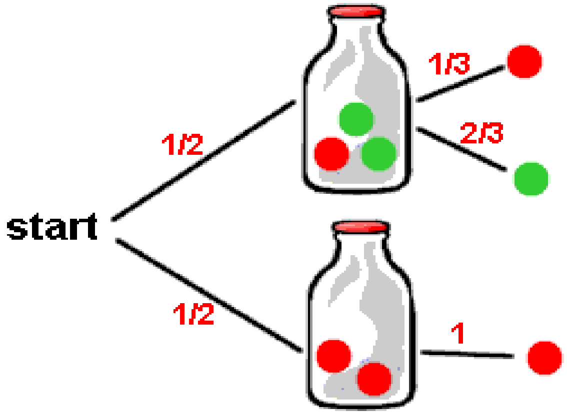
$$\bigcup_{k=1}^{n} F_k = S, \text{ where } F_k F_j = \phi, k \neq j.$$

- Then the total probability rule can be generalized as follows, assuming the above,

$$P(E) = \sum_{\mathrm{k}=1}^{n} P(E \mid F_{\mathrm{k}})\, P(F_{\mathrm{k}}).$$

Red ball from a bottle: revisited

- There are two identical bottles. One contains 2 green balls and 1 red ball, the other bottle contains 2 red balls. A bottle is selected at random and a ball is drawn. The ball is red. What are the chances that it was drawn from the first bottle?



Solution

- To find is *P(I | Red)*. By definition,

$$P(\mathrm{I} \mid \mathrm{Red}) = \frac{P(\mathrm{I\,Red})}{P(\mathrm{Red})} = \frac{P(\mathrm{Red} \mid \mathrm{I})P(\mathrm{I})}{P(\mathrm{Red})}$$

- Hence,

$$P(\mathrm{I} \mid \mathrm{Red}) = \frac{\frac{1}{3} \cdot \frac{1}{2}}{\frac{1}{3} \cdot \frac{1}{2} + 1 \cdot \frac{1}{2}} = \frac{\frac{1}{6}}{\frac{2}{3}} = \frac{1}{4}$$

Analysis

- In the example, the event that a red ball was drawn from a bottle has affected the chances that it was a particular bottle.
- Originally, both bottles were equally likely, i.e., $P(\mathrm{I}) = P(\mathrm{II}) = 0.5$.
- With a red ball drawn, it is three times more likely that the second bottle was selected (i.e., the odds are 3), as the respective conditional probability is 0.75 vs. 0.25 for Bottle 1.
- This was an example of the ***Bayes' Formula***. Also called Bayes' Rule or Bayes' Theorem.

Bayes' Formula: General Case.

Thomas Bayes
(1702-1761)

- Suppose that S is represented as the union of n disjoint events F_k.
- Then

$$P(F_i|E) = \frac{P(E|F_i)\,P(F_i)}{\sum_{k=1}^{n} P(E|F_k)P(F_k)},$$

where

$$\bigcup_{k=1}^{n} F_k = S, \text{ and } F_k F_j = \phi,\ j \neq k$$

Example

- In semiconductor manufacturing, a chip is subject to high, medium or low contamination levels. The probability of chip failure is 0.1, if contamination level is high, 0.01, if it is medium, and 0.001, if it is low. 20% of chips are subject to high and 30% to medium contamination levels. If a chip causes a product failure, find the probability that the chip was subject to high contamination levels.
- Hint: use the Bayes' formula.

Solution to "contaminated chip questions"

- Let L, M and H stand for the events that contamination levels were low, medium and high, respectively. Then

$$P(L) = 0.5,\ P(M) = 0.3,\ P(H) = 0.2$$

- Let F be the event that a chip causes a failure. Then

$$P(F|L) = 0.001,\ P(F|M) = 0.01,\ P(F|H) = 0.1$$

- Then the total probability rule implies

$$P(F) = 0.0005 + 0.003 + 0.02 = 0.0235.$$

Solution to the second part

The objective is to find $P(H|F)$.

By the Bayes' Formula,

$$P(H \mid F) = \frac{P(F \mid H)P(H)}{P(F)} = \frac{0.02}{0.0235} \approx 0.851$$

Example

- Duff Cola Company recently received several complaints that their bottles are under-filled.
- A complaint was received today but the production manager is unable to identify which of the two Springfield plants (A or B) filled this bottle. The following is known.

Plant	% of Total Production	% of Under-filled Bottles
A	55	3
B	45	4

- What is the probability that the under-filled bottle came from plant A?

Solution

- Let A and B stand for the events that the bottle was filled at plant A and B, respectively. Then

$$P(A) = 0.55,\ P(B) = 0.45$$

- Let U be the event that the bottle is under-filled. Then

$$P(U|A) = 0.03,\ P(U|B) = 0.04$$

Then the total probability rule implies

$$\begin{aligned} P(U) &= P(U|A)P(A) + P(U|B)P(B) \\ &= 0.03(0.55)+0.04(0.45) \\ &= 0.0345. \end{aligned}$$

Solution, cont.

We want to find *P(A|U)*. By the Bayes' Formula,

$$P(A\,|\,U) = \frac{P(U\,|\,A)P(A)}{P(U)} = \frac{0.03(0.55)}{0.0345} \approx 0.478$$

Therefore, although there is a 55% chance that the bottle was filled at plant *A*, if it is under-filled, there is a 47.8% chance that it was filled at plant *A*.

Example - Revisited

- One card is drawn from a standard deck, looked at, replaced, and then the second card is drawn. Find the probability that the first card is a king, and the second one is not.
- Solution: The sample space S is different this time, namely, $N(S)=52x52$.
- The number of outcomes in E is – by the multiplication rule – $N(E) = 4x48$ (4 kings for the first card and 48 not-kings for the second one).
- Hence, $P(E) = (4x48)/(52x52)$.

A closer look

$$P(E) = \frac{4 \cdot 48}{52 \cdot 52} = \frac{4}{52} \cdot \frac{48}{52}$$

$$E = K_1 N_2$$

$$P(K_1) = \frac{4}{52}, \; so$$

$$P(K_1 N_2) = P(K_1) \cdot P(N_2)$$

Independent Events

- In such a case it is natural to call events N_2 and K_1 ***independent***.
- In general, two events, *E* and *F*, are called ***independent*** if

$$P(EF) = P(E) \cdot P(F)$$

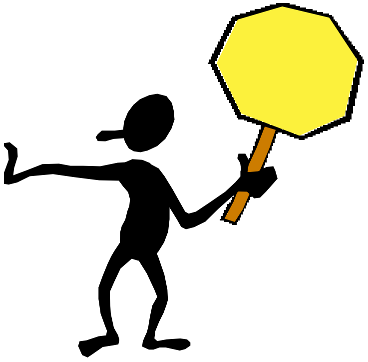
$$or$$

$$P(E \mid F) = P(E)$$

Comments

- In practice, the conditions of an experiment allow one to conclude that certain events are independent.
- In this case the definition of independence is used to compute required probabilities.
- One typical example is independently working relays in an electric circuit.
- Or, transmission of bits via a digital communication channel.

Warning



- Do not be confused between *mutually exclusive* and *independent* events!
- If two events, *E* and *F*, are mutually exclusive, it means that if E occurs, then F cannot occur and vice-versa.
- If two events, E and F, are independent, then it means that occurrence of F does not depend on occurrence of E and vice versa.
- For mutually exclusive events, we add probabilities and for independent events we multiply.
- If E and F are independent and mutually exclusive then,

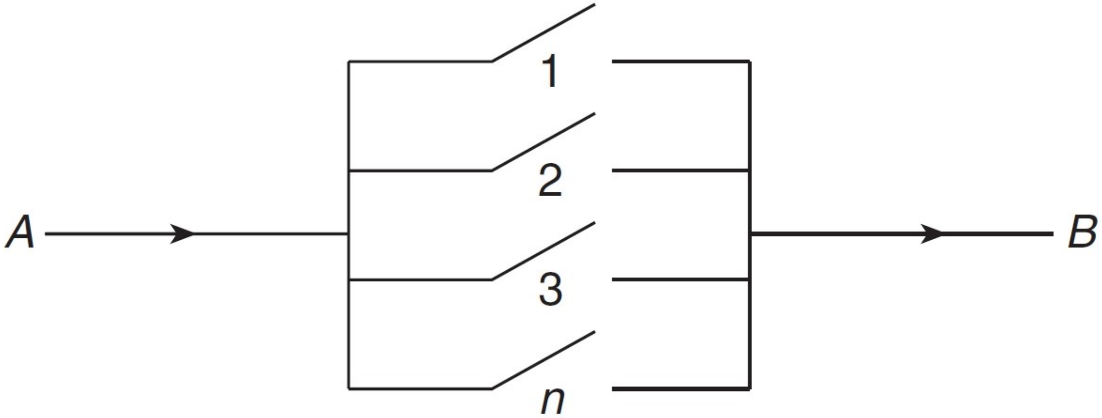
$$0 = P(\phi) = P(EF) = P(E) \cdot P(F)$$

Independence (more than 2 events)

- The notion of independence can be easily extended to a finite number of events.
- Three events, *E,F,G*, are independent if each two of them are independent (i.e., pairwise independent), and, in addition,

$$P(EFG) = P(E) \cdot P(F) \cdot P(G)$$

Example



- *A* system is composed of n separate components.
- It is ***parallel*** if the system functions when at least one component does.
- It is also assumed that the operation of one component does not affect the other ones.
- Suppose that the probability that the k-th component functions is p_k. Thus the probability that the k-th component fails is $1 - p_k$.
- *Find the probability that the system functions.*

Solution

- Let A be the event that the system functions.
- The system is *parallel*, hence

$$A = A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_n = \bigcup_{k=1}^{n} A_k$$

where A_k is the event that the k-th component functions.

Solution

- By De Morgan Laws, $A^c = \left(\bigcup_{k=1}^{n} A_k \right)^c = A_1^c A_2^c A_3^c \cdots A_n^c$
- Independence implies that

$$P(A^c) = \prod_{k=1}^{n} P(A_k^c) = \prod_{k=1}^{n} (1 - p_k)$$

- Hence,

$$P(A) = 1 - P(A^c) = 1 - \prod_{k=1}^{n} (1 - p_k)$$

Food for Thought – Example 2.1

- In answering a question on a multiple-choice test, a student either knows the answer or she guesses. Let p=0.6 be the probability that she knows the answer and 1−p = 0.4 be the probability that she guesses.
- Assume that a student who guesses at the answer will be correct with probability 1/m, where m is the number of multiple-choice alternatives.
- What is the conditional probability that a student knew the answer to a question with m=4 given that she answered it correctly?

Food for Thought – Example 2.2

- You ask your neighbour to water a sickly plant while you are on vacation. Without water it will die with probability .8; with water it will die with probability .15. You are 90 percent certain that your neighbour will remember to water the plant.
  - What is the probability that the plant will be alive when you return?
  - If it is dead, what is the probability your neighbour forgot to water it?

More on Independence - next class

- Further examples of independent events are going to motivate the use of binomial coefficients to deal with large numbers of outcomes, as we will see.
- Further, next week we are going to discuss how to count in random setting.
- So next class' topic is

RANDOM VARIABLES

References/Resources Used

- Lecture Slides for MATH 401 of Dr. Oleksiy Us, Department of Mathematics, German University of Cairo. [PPT]

Food for Thought – Example 2.1

- In answering a question on a multiple-choice test, a student either knows the answer or she guesses. Let $p=0.6$ be the probability that she knows the answer and $1-p = 0.4$ be the probability that she guesses.
- Assume that a student who guesses at the answer will be correct with probability $1/m$, where m is the number of multiple-choice alternatives.
- What is the conditional probability that a student knew the answer to a question with $m=4$ given that she answered it correctly?

Solution for Example 2.1

- Let C and K denote, respectively, the event that the student answers the question correctly and the event that she actually knows the answer. We want to compute P(K|C). We have

$$P(K|C) = \frac{P(KC)}{P(C)}$$

- Now

$$\begin{aligned} P(KC) &= P(K)P(C \mid K) \\ &= (0.6)(1) \\ &= 0.6 \end{aligned}$$

Solution for Example 2.1

- To compute the probability that the student answers correctly, we condition on whether or not she knows the answer. That is,

$$P(C) = P(C \mid K)P(K) + P(C \mid K^c)P(K^c)$$
$$= (1)(0.6) + (0.25)(0.4) = 0.7$$

- Hence, the desired probability that a student knew the answer to a question she correctly answered is given by

$$P(K|C) = \frac{P(KC)}{P(C)} = \frac{0.6}{0.7} = 0.857$$

Food for Thought – Example 2.2

- You ask your neighbour to water a sickly plant while you are on vacation. Without water it will die with probability .8; with water it will die with probability .15. You are 90 percent certain that your neighbour will remember to water the plant.
  - What is the probability that the plant will be alive when you return?
  - If it is dead, what is the probability your neighbour forgot to water it?

Solution for Example 2.2

- Let A and W denote, respectively, the event that the plant is alive when you get back and the event that neighbour watered the plant in your absence. First, we want to compute P(A). We have P(W) = 0.9, P(A|W) = 0.85, and P(A|W^c) = 0.2. By the *total probability rule* the requested probability is

$$
\begin{aligned}
P(A) &= P(A\,|\,W)P(W) + P(A\,|\,W^c)P(W^c) \\
&= (0.85)(0.9) + (0.2)(0.1) \\
&= 0.785
\end{aligned}
$$

Solution for Example 2.2

- The probability that if it is dead, then your neighbour forgot to water it is given by $P(W^c|A^c)$. We have

$$P(W^c|A^c) = \frac{P(W^c A^c)}{P(A^c)} = \frac{P(W^c A^c)}{\left(1 - P(A)\right)}$$

Also, $P(W^c A^c) = P(A^c W^c) = P(A^c|W^c)P(W^c) = (0.8)(0.1)$

Finally,

$$P(W^c|A^c) = \frac{P(W^c A^c)}{\left(1 - P(A)\right)} = \frac{(0.8)(0.1)}{(1 - 0.785)} = \frac{0.08}{0.215} \approx 0.372$$

Food for thought - Example 2.3

- In semiconductor manufacturing, a chip is subject to high, medium or low contamination levels. The probability of chip failure is 0.1, if contamination level is high, 0.01, if it is medium, and 0.001, if it is low. 20% of chips are subject to high and 30% to medium contamination levels. If a chip causes a product failure, find the probability that the chip was subject to high contamination levels.
- Hint: use the Bayes' formula.

Solution to "contaminated chip questions"

- Let L, M and H stand for the events that contamination levels were low, medium and high, respectively. Then

$$P(L) = 0.5,\ P(M) = 0.3,\ P(H) = 0.2$$

- Let F be the event that a chip causes a failure. Then

$$P(F|L) = 0.001,\ P(F|M) = 0.01,\ P(F|H) = 0.1$$

- Then the total probability rule implies

$$P(F) = 0.0005 + 0.003 + 0.02 = 0.0235.$$

Solution to the second part

The objective is to find $P(H|F)$.

By the Bayes' Formula,

$$P(H \mid F) = \frac{P(F \mid H)P(H)}{P(F)} = \frac{0.02}{0.0235} \approx 0.851$$