Sample Space and Sample Points (Continued)

1.1.1 Sample Space (Continued)

- **Sample Space:** The set of all possible outcomes of a statistical experiment is called the **sample space** and it is represented by the symbol S.

One important point here one should pay attention is that the "Sample Space" depends on not only how the experiment is carried out, but also the "problem of interest". Also refer to page 7 for an example.

Examples

1. Consider an experiment of tossing a die.
   - If we are interested in the number that shows on the top face, then the sample space would be
     $S = \{1, 2, 3, 4, 5, 6\}$.
   - If we are interested only in whether the number is even or odd, then the sample space is simply
     S = {even, odd}.

Examples (Continued)

2. Consider an experiment of tossing two dice.
   – If we are interested in the numbers that show on the top faces, then the sample space would be

$$S = \{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,1), (2,2), \cdots\cdots, (6,5), (6,6)\}.$$

If the dice are labelled, i.e., whether the first die shows 1 or the second shows 1 matters, the sample space is the one given in the slides; each observation in this sample space is equally likely to appear. If the two dice are not labelled, the sample space is

$$\begin{aligned} S = \{&\{1,1\}, \{1,2\}, \{1,3\}, \{1,4\}, \{1,5\} \\ &\{2,2\}, \{2,3\}, \{2,4\}, \{2,5\}, \{2,6\} \\ &\{3,3\}, \{3,4\}, \{3,5\}, \{3,6\} \\ &\{4,4\}, \{4,5\}, \{4,6\} \\ &\{5,5\}, \{5,6\} \\ &\{6,6\}\} \end{aligned}$$

But now, take note that each observation in the sample space is not equally likely; in other words, their chance of appearance are not the same. In particular, "{1,1}" is less likely to appear than "{1,2}".

1.1.3 Events

An **event** is a subset of a sample space.

Examples

1. (a) $S = \{1, 2, 3, 4, 5, 6\}$.
   An event that an odd number occurs $= \{1, 3, 5\}$
   An event that a number greater than 4 occurs $= \{5, 6\}$

   (b) $S = \{\text{even}, \text{odd}\}$.
   An event that an odd number occurs $= \{\text{odd}\}$

ST2334 Probability and Statistics CYM Basic Concepts of Probability 1-15

In principle, any arbitrary subset of the sample space can be an event. So, for a sample space with n number of possible sample points, there are 2^n number of possible events (we shall be able to verify this rigorously after we have learned "combinations"), including the empty event (null event), denoted by Ø by convention. Here empty event is the event which contains no sample point. So, if S={1,2,3,4,5,6}, the number of possible event is $2^6 = 64$, which is quite a big number, and it is not feasible for us to list them all. In this slide, it gives some frequently used events in practice.

Here is one take home question to practice: if S = {1,2,3} so that we have 8 possible events, please try to write down all the possible events relating to this sample space.

For this case, the list of events are $\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}$.

Page 29 of Chapter1 Part 1

Examples (Continued)

$S = \{1, 2, 3, 4, 5, 6\}$. $A = \{1, 2, 3\}$, $B = \{1, 3, 5\}$ and $C = \{2, 4, 6\}$.

Then

- $A' = \{4, 5, 6\}$
- $B' = \{2, 4, 6\} = C$
- $A' \cap B' = \{4, 5, 6\} \cap \{2, 4, 6\} = \{4, 6\}$
- $(A \cup B)' = \{1, 2, 3, 5\}' = \{4, 6\}$

Notice that both $A' \cap B'$ and $(A \cup B)'$ equal $\{4, 6\}$ in this example. Is it true that $A' \cap B' = (A \cup B)'$ in general?

ST2334 Probability and Statistics CYM Basic Concepts of Probability 1-29

$(A \cup B)' = A' \cap B'$ is in fact the special case of the De Morgan's Law. So you may read this caculation jointly with page 36:

1.2.7 De Morgan's Law

For any n events $A_1, A_2, \cdots, A_n$,

1.
$$(A_1 \cup A_2 \cup \cdots \cup A_n)' = A_1' \cap A_2' \cap \cdots \cap A_n'$$

2.
$$(A_1 \cap A_2 \cap \cdots \cap A_n)' = A_1' \cup A_2' \cup \cdots \cup A_n'$$

ST2334 Probability and Statistics CYM Basic Concepts of Probability 1-36

Also, you can refer to page 34 and 35, points 5 and 6.

Page 34&35 of Chapter1 Part 1

1.2.6 Some Basic Properties of Operations of Events

1. $A \cap A' = \emptyset$.
2. $A \cap \emptyset = \emptyset$.
3. $A \cup A' = S$.
4. $(A')' = A$
5. $(A \cap B)' = A' \cup B'$

ST2334 Probability and Statistics CYM Basic Concepts of Probability 1-34

Some Basic Properties of Operations of Events (Continued)

6. $(A \cup B)' = A' \cap B'$
7. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
8. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
9. $A \cup B = A \cup (B \cap A')$
10. $A = (A \cap B) \cup (A \cap B')$

ST2334 Probability and Statistics CYM Basic Concepts of Probability 1-35

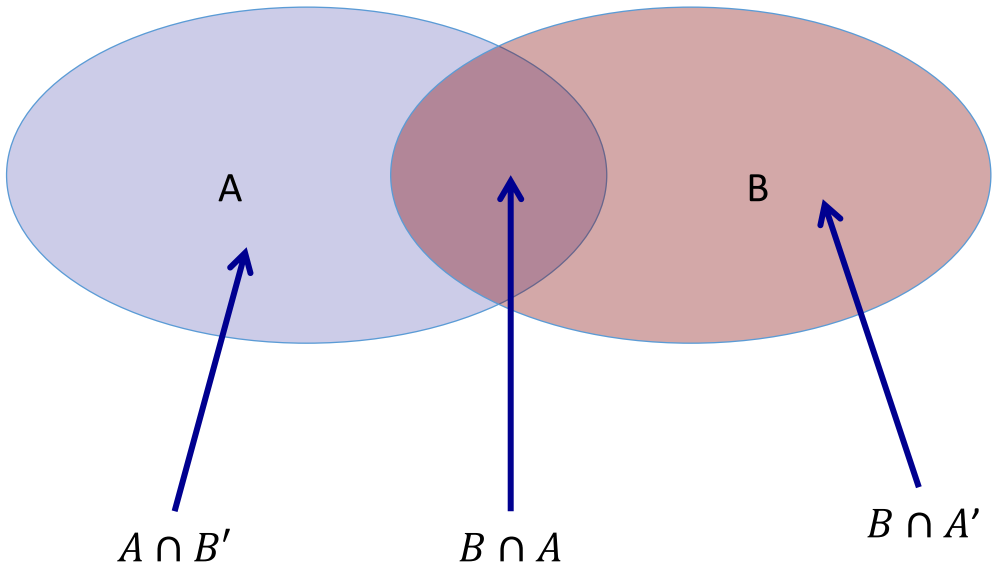
- These are very useful formulae.
- Besides, you may also add the formulae discussed in pages 27 and 28:

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

$$A \cup B \cup C = (A \cup B) \cup C = A \cup (B \cup C)$$

- Also take note that $A \cap (B \cup C) \neq (A \cap B) \cup C$; in other words, when the operations are not of the same type, which operation being done first matters.

- Property 9 in the above slides essentially gives one mutually exclusive partition for $A \cup B$, in other words, besides $A \cup B = A \cup (B \cap A')$, we should also observe $A \cap (B \cap A') = \emptyset$. So, is there any other mutually exclusive partition on $A \cup B$? With the help of the venn diagram, one can indeed identify at least two other possible partitions. Here we give one example of such a partition: $A \cap B, A \cap B', B \cap A'$; the other is left as an excise.
- Property 10 in the above slides have given a mutually exclusive partition for A, which means besides $A = (A \cap B) \cup (A \cap B')$, we also have $(A \cap B) \cap (A \cap B') = \emptyset$. Can you also come up with a similar partition for B?



Page 62 of Chapter1 Part 1

Example 7

In how many ways can 4 boys and 5 girls sit in a row if the boys and girls must alternate?

ST2334 Probability and Statistics CYM Basic Concepts of Probability 1-62

If 5 boys and 5 girls are to sit in a row, such that boys and girls are sitting alternatively, how many ways can we make the arrangement?

There are two sitting plans:

Plan 1: BGBGBGBGBG

Plan 2: GBGBGBGBGB

Therefore, the total number of ways is the number of ways for Plan 1 plus the number of ways for Plan 2.

For Plan 1, the number of ways is $5 \times 5 \times 4 \times 4 \times 3 \times 3 \times 2 \times 2 \times 1 \times 1 = 5! \times 5!$; and likewise, for Plan 2, the number of ways is $5! \times 5!$ as well. So the total number of ways is $2 \times 5! \times 5! = 28800$.

Page 62 of Chapter1 Part 1

1.3.2 Addition Principle

- Suppose that a procedure, designated by 1 can be performed in n_1 ways.
- Assume that a procedure, designated by 2 can be performed in n_2 ways.
- Suppose furthermore that it is **NOT possible** that both procedures 1 and 2 are **performed together**.
- Then the number of ways in which we can perform 1 **or** 2 is

$$n_1 + n_2$$

ways.

ST2334 Probability and Statistics CYM Basic Concepts of Probability 1-64

If we are doing parallel counting, “addition principle” takes effect; if we are doing sequential counting, “multiplication rule” applies. The example in the last page in fact has jointly used both principles. In particular, “Plan 1” and “Plan 2” parallel, so “addition principle” should be used; therefore, we add the number ways based on “Plan 1” and those based on “Plan 2”. Then, when computing the number of ways for each Plan, Plan 1 say, “multiplication principle” has been applied, we assume that kids are sitting from left to right sequentially.

Binomial Coefficient

- The quantity $\binom{n}{r}$ is called a **binomial coefficient** because it is the coefficient of the term $a^r b^{(n-r)}$ in the binomial expansion of $(a+b)^n$.
- It can be verified that the following hold:
  1. $\binom{n}{r} = \binom{n}{n-r}$ for $r = 0, 1, \cdots, n$
  2. $\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}$ for $1 \le r \le n$
  3. $\binom{n}{r} = 0$ for $r < 0$ or $r > n$

ST2334 Probability and Statistics CYM Basic Concepts of Probability 1-107

The first statement in this page means:

$(a+b)^n = \binom{n}{n}a^n + \binom{n}{n-1}a^{n-1}b + \binom{n}{n-2}a^{n-2}b^2 + \cdots + \binom{n}{1}ab^{n-1} + \binom{n}{0}b^n$. (1)

A direct consequence of this formula is, by setting $a = b = 1$,

$$2^n = \binom{n}{n} + \binom{n}{n-1} + \binom{n}{n-2} + \cdots + \binom{n}{1} + \binom{n}{0}.$$

Note: This last formula is able to be used to verify the statement we claimed in the 3rd page of this document: "for a sample space with n number of possible sample points, there are 2^n number of possible events".

An approach to achieve the expansion in (1) is as follows:

- Let's think $(a+b)^n = (a+b)(a+b)\cdots(a+b) = Factor1 \times Factor2 \times \cdots \times Factorn$, i.e., the multiplication of n number of "$(a+b)$"s. When expanding this formula directly with algebra, we shall get 2^n expanded terms: $(a+b)(a+b)\cdots(a+b) = E1 + E2 + \cdots + E2^n$. Here Ei is a product of n factors, i.e., $Ei = T1 \times T2 \times \cdots \times Tn$ with $T1(= a\ or\ b)$ from $Factor1$, $T2(= a\ or\ b)$ from $Factor2$, and so on.
- Now, we can count that among all these $Ei's$:

- the number of Ei's that are equal to a^n should be $\binom{n}{n}$: all the corresponding $T1, T2, \cdots, Tn$ must be a. So the coefficient for a^n is $\binom{n}{n}$.
- the number of Ei's that are equal to $a^{n-1}b$ should be $\binom{n}{n-1}$: out of all the $T1, T2, \cdots, Tn$, $n-1$ of them must be a, and the rest must be b. So the coefficient for $a^{n-1}b$ is $\binom{n}{n-1}$.
- the number of Ei's that are equal to $a^{n-2}b^2$ should be $\binom{n}{n-2}$: out of all the $T1, T2, \cdots, Tn$, $n-2$ of them must be a, and the rest must be b. So the coefficient for $a^{n-2}b^2$ is $\binom{n}{n-2}$.
- …..
- the number of Ei's that are equal to b^n should be $\binom{n}{0}$: out of all the $T1, T2, \cdots, Tn$, 0 of them must be a, and the rest must be b. So the coefficient for b^n is $\binom{n}{0}$.

Pages 125&128 of Chapter1 Part 1

1.4.2 Relative Frequency

- Alternatively, probability can be derived based on the relative frequency
- Suppose we repeat the experiment E for n times and let A be an event associated with E.
- We let n_A be the number of times that the event A has occurred among the n repetitions respectively.
- Then $f_A = \frac{n_A}{n}$ is called the ***relative frequency*** of the event A in the n repetitions of E.

1.4.3 Axioms of Probability

- Consider an experiment whose sample space is S.
- The objective of probability is to assign to each event A, a number $\Pr(A)$, called the **probability** of the event A, which will give a precise measure of the chance that A will occur.
- Consider the collection of all events and denote it by $\mathcal{P}$.
- For each event A of the sample space S we assume that a number $\Pr(A)$, which is called the **probability** of the event A, is defined and satisfies the following three axioms:

Relative frequency and Probability are very close, but different terminologies. One can view the former as the practical explanation of the latter in some sense. More rigorously, we can assert that $f_A = n_A/n$ approaches $\Pr(A)$ when n approaches infinity, or described more mathematically:

$$\Pr(A) = \lim_{n\to\infty} f_A$$

Axioms of Probability (Continued)

Axiom 1: $0 \leq \Pr(A) \leq 1$.

Axiom 2: $\Pr(S) = 1$.

Axiom 3: If $A_1, A_2, \cdots$ are **mutually exclusive** (disjoint) events (that is, $A_i \cap A_j = \emptyset$ when $i \neq j$), then

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

In particular, if A and B are two mutually exclusive events then $\Pr(A \cup B) = \Pr(A) + \Pr(B)$.

These Axioms, in many literatures, are used as the definition of probability.

One should take note that

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

and that in page 127: $f_{A\cup B} = f_A + f_B$ are correct ONLY IF the events for taking the unions are MUTUALLY EXCLUSIVE!

Page 137 of Chapter 1 Part 2

1.5 Basic Properties of Probability

1.5.1 Some Basic properties of probability

1. $\Pr(\emptyset) = 0$.

   Take $A_1 = \emptyset, A_2 = \emptyset, A_3 = \emptyset, \cdots\cdots$

   then $A_1 \cap A_2 \cap A_3 \cap \cdots\cdots = \emptyset$ and $A_1 \cup A_2 \cup A_3 \cup \cdots\cdots = \emptyset$.

   Hence, by Axiom 3

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

   which can hold only if $\Pr(\emptyset) = 0$.

This can be derived slightly simpler (but the same idea): set $A_1 = \emptyset$ and $A_2 = \emptyset$. Then

$\Pr(\emptyset) = \Pr(A_1 \cup A_2) = \Pr(A_1) + \Pr(A_2) = \Pr(\emptyset) + \Pr(\emptyset)$, which leads to $\Pr(\emptyset) = 0$.

Note that the inverse is not necessarily true: $\Pr(A) = 0$ does not imply $A = \emptyset$. A can contain sample points that has zero probability. For example, consider the outcome of rolling a die and $A = \{7,8,9\}$, then $\Pr(A) = 0$, but A is not empty. Here is another example, randomly select a whole number, i.e., all whole numbers are equally likely to be chosen, then S is the set of whole numbers. $A = \{1,2,3,4,5,6\}$, then $\Pr(A) = 0$; but A is not empty.

Page 145 of Chapter 1 Part 2

Basic Properties of Probability (Continued)

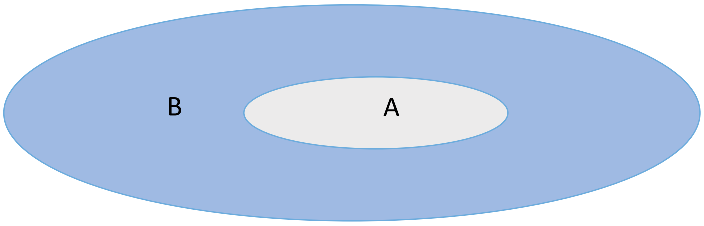
7. If $A \subset B$, then $\Pr(A) \le \Pr(B)$.

Since $B = (B \cap A) \cup (B \cap A')$ and $B \cap A = A$,
so $B = A \cup (B \cap A')$ and $A \cap (B \cap A') = \emptyset$.
Therefore

$$\Pr(B) = \Pr(A) + \Pr(B \cap A') \ge \Pr(A).$$

The result follows by noting that $\Pr(B \cap A') \ge 0$.

This derivation can be illustrated by the following diagram:



Here, B is includes both the blue and gray areas; A is the gray area, $B \cap A'$ is the blue area. So we immediately observe $B = A \cup (B \cap A')$, and $\emptyset = A \cap (B \cap A')$. Therefore,

$$\Pr(B) = \Pr(A) + \Pr\,(B \cap A') \ge \Pr\,(A)\,.$$

Based on this derivation, we can answer more involved quesitons: Can $\Pr(B) = \Pr(A)$? If yes, when? When can we conclude $\Pr(B) > \Pr(A)$, i.e., we can remove the possibility of "="? The answer is fully due to the probability event $B \cap A'$ and its probability.

Sometimes, it is not so obvious to conclude that two events are multually exclusive, then we may need to justify with some mathematical arguments. To conclude two events, namely C and D, are mutually exclusive, we need to

- Assume an arbitrary $x \in C$, then verify $x \notin D$;
- And assume an arbitrary $x \in D$, then verify $x \notin C$.

Therefore, we can check that A and $B \cap A'$are mutually exclusive by

- Assume an arbitrary $x \in A$, then clearly $x \notin A'$ based on the definition of A'; therefore we conclude $x \notin B \cap A'$ based on the defintion of intersection.
- Assume an arbirary $x \in B \cap A'$, then we must have $x \in A'$, so $x \notin A$.

Note that we can extend such arguments to check other relationships of sets.

To check $A = B$, we need to

- Assume an arbitrary $x \in A$, then verify $x \in B$;
- And assume an arbitrary $x \in B$, then verify $x \in A$.

To check $A \subset B$, i.e., A is a subset of B, we only need to assume arbitrary $x \in A$, then verify $x \in B$.

To check that $A \not\subset B$, we only need to find an $x \in A$, and verify that $x \notin B$.

Solution to Example 6

- The probability that n persons all have different birthdays **from you** is $\left(\frac{364}{365}\right)^n$.
- So we need n such that $1-\left(\frac{364}{365}\right)^n \geq 0.5$.
- Solving, we obtain
$$n \geq \frac{\log(0.5)}{\log(364/365)} = 252.7.$$
- We need at least 253 people (excluding yourself).

ST2334 Probability and Statistics CYM Basic Concepts of Probability 1-162

The "blue formula" in this lecture slide can be derived in the same manner as the "Birthday problem":

- In total, there are 365 possible choices as the birthday of each person. So if there are n persons, the total number of possibilities of their birthdays would be 365^n.
- If we require that no body would be of the same birthday as you. The number of possible choices as the birthday of each person would be 364. So there are 364^n possibilities.

This leads to the "blue formula" in the slide.

To solve the inequality $1-\left(\frac{364}{365}\right)^n \geq 0.5$,

- Moving terms around, we get $\left(\frac{364}{365}\right)^n \leq 0.5$;
- Taking log on both sides of the inequality leads to $n\log\left(\frac{364}{365}\right) \leq \log(0.5)$, since for any positive number x, y, $x \leq y$ if an only if $\log(x) \leq \log(y)$.
- Note that $\log\left(\frac{364}{365}\right) < 0$, as $\frac{364}{365} < 1$. Then devide a negative number on both sides of the inequality, the direction of the inequality needs to be changed. Therefore $n \geq \log(0.5)/\log(364/365)$.

1.5.2 Sample Spaces Having Finite Outcomes

Consider the sample space S which contains a finite number of k outcomes. That is,

$$S = \{a_1, a_2, \cdots, a_k\}$$

Let $\Pr(a_i) = p_i$ be the probability of $\{a_i\}$ and

(1) $0 \leq p_i \leq 1$, for $i = 1, 2, \cdots, k$.

(2) $p_1 + p_2 + \cdots + p_k = 1$.

$p_1 + p_2 + \cdots + p_k = 1$ can be derived as follows. Define probability events: $A_1 = \{a_1\}, A_2 = \{a_2\}, \cdots, A_k = \{a_k\}$. Then clearly all these events are mutually exclusive as each only carries one sample point, which are distinct to each other, and $S = A_1 \cup A_2 \cup \cdots \cup A_k$. Therefore

$$1 = \Pr(S) = \Pr(A_1 \cup A_2 \cup \cdots \cup A_k) = \Pr(A_1) + \Pr(A_2) + \cdots + \Pr(A_k)$$

$$= \Pr(a_1) + \Pr(a_2) + \cdots + \Pr(a_k) = p_1 + p_2 + \cdots + p_k.$$

Example 3

- A box contains 50 bolts and 150 nuts.
- Half of the bolts and half of the nuts are rusted.
- If one item is chosen at random, what is the probability that it is rusted or is a bolt?

Solution

- Let A = {the item is rusted} and B = {the item is a bolt}.
- Then

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$$
$$= 100/200 + 50/200 - 25/200 = 5/8.$$

This example can also be worked out directly by the formula given on page 172:

$$\Pr(A) = \frac{Number\ of\ Sample\ Points\ in\ A}{Number\ of\ Sample\ points\ in\ S}$$

Here, we define $A = \{the\ item\ is\ rusted\ or\ a\ bolt\}$, which is the collection of all the bolts and the rusted nuts; so in total, we have 50+75 = 125 sample points in A; we have 200 sample points in S. Therefore $\Pr(A) = \frac{125}{200} = 5/8$, which is the same as the answer in the slide.

Page 174 of Chapter 1 Part 2

Example 5

- If 2 balls are "randomly drawn" from an urn containing 6 white and 5 black balls,
- what is the probability that one of the drawn balls is white and the other black?

ST2334 Probability and Statistics CYM Basic Concepts of Probability 1-176

Page 177 gives the way to solve this problem using combinations, where which ball is drawn first does not matter. To solve this problem but considering that the order of drawing the balls matter, we can think this way:

- The total number of ways to draw is: for the first draw, there are 11 possibilities, and 10 for the second; so based on the multiplication principle, we have $(11)(10) = 110$, which is the number of sample points in the sample space.
- To get a white ball and a black ball, there are two parallell possibilities: WB (the first ball is white, and the second is black) or BW. To get WB, there are (6)(5)=30 possibilies: draw the first ball from those 6 white balls, and draw the second from the 5 black balls. Likewise, to get BW, there are (5)(6) = 30 possibilities. So in total, we have 60 possible ways to get one white ball and one black ball.

As a consequence, the probability of getting a white ball and a black ball is 60/110 = 6/11.

Page 186&188 of Chapter 1 Part 2

Example 9

- In a class of 200 students, 108 study economics, 138 study chemistry and 70 study both chemistry and economics.
- If a student is selected at random, what is the probability that the student
  - (a) takes economics **or** chemistry;
  - (b) doesn't take **neither** of these subjects;
  - (c) takes chemistry **but not** economics.

Part (b) needs to be corrected to be "take neither of these subjects".

Solution to Example 9 (Continued)

(c) Since $C = (E' \cap C) \cup (E \cap C)$ and $(E' \cap C) \cap (E \cap C) = \emptyset$, therefore

$$\Pr(C) = \Pr(E' \cap C) + \Pr(E \cap C).$$

Hence $\Pr(E' \cap C) = \Pr(C) - \Pr(E \cap C)$

$$= 138/200 - 70/200$$

$$= 0.34.$$

$C = (E' \cap C) \cup (E \cap C)$ can be derived using formulae of sets computations:

$$C = S \cap C = (E' \cup E) \cap C = (E' \cap C) \cup (E \cap C).$$

Besides, think about whether there is another way to solve (c).

An Illustrative Example (Continued)

- Hence the conditional probability of A given B can be obtained by
$$\Pr(A|B) = \frac{\#(A \cap B)}{\#(B)}$$
- If we divide both the numerator and denominator by $\#(S)$, then
$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}.$$
- Similarly,
$$\Pr(B|A) = \frac{\Pr(A \cap B)}{\Pr(A)}.$$

- From this page, we can clearly observe $Pr(A|B) \neq Pr(B|A)$, unless $Pr(B) = Pr(A)$. Also take note that this two conditional probabilies have totally different meanings. Pr(A|B) is confining the probability space to B (namely B is given), whereaa P(B|A) is confining probability space to A (namely A is given).
- We also have $Pr(A|B') \neq 1 - Pr(A|B)$ in general. But $\Pr(A'|B) = 1 - P(A|B)$.
- The conclusion $\Pr(A|B) = \frac{\#(A \cap B)}{\#B}$ is applicable only when the sample points in the sample space is equally likely. The example on page 198 gives an example that this formula is not applicable:

Example 2

- Roll an unbalanced die.
  (The sample space is $S = \{1, 2, 3, 4, 5, 6\}$.)
- Suppose that the respective probabilities are $1/12$, $1/12$, $1/6$, $1/6$, $1/6$, and $1/3$.
  Note: The sum of the probabilities equals 1
  (a) If the number obtained is even, what is the probability that it is a 6?
  (b) What is the probability that the number obtained is a perfect square number given that a number greater than 3 has obtained?

The following page gives another example that it is essential to ensure that the sample points are equally likely in order to apply this formula; refer to the corresponding lecture video for this example for details.

Example 3

- A couple has 2 children.
- What is the probability that both are boys if it is known that they have at least one son?