Chapter 3: Elements Of Probability

ST2334 Probability and Statistics[1]
(Academic Year 2014/15, Semester 1)

Department of Statistics & Applied Probability (DSAP)
National University of Singapore (NUS)[1]

Outline

Introduction I

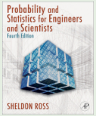
Learning Outcomes

Questions to Address: ♦ Understanding algebra of events ♦ How to compute/interpret probability ♦ Various properties of probability ♦ How to count ♦ Difference between combination & permutation ♦ What conditional probability is ♦ What independence is

Concept & Terminology: ♦ experiment ♦ sample space ♦ set theory ♦ event ♦ complement/union/intersection ♦ null/mutually exclusive/disjoint/exhaustive events ♦ Venn diagram ♦ commutative/associative/distributive/DeMorgan's laws ♦ axioms of probability ♦ addition/multiplication law ♦ equiprobable outcomes ♦ basic principle of counting ♦ permutation ♦ combination ♦ binomial coefficient ♦ conditional probability ♦ law of total probability ♦ tree diagram ♦ Bayes' formula ♦ independence ♦ pairwise/mutual independence

Introduction II

Mandatory Reading

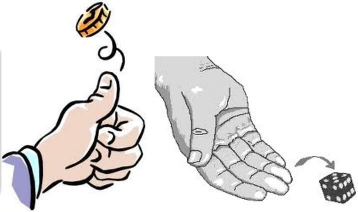


✏ Section 3.1 – Section 3.8

Sample Spaces

Definition

An *experiment* is any action or process whose outcome is subject to uncertainty/randomness



- When an *experiment is performed/conducted*, *ONLY one* of all possible outcomes would *occur* or would be *observed/realized*
- Assume that the *set/collection of all possible outcomes* is known

Definition

The *sample space* of an experiment, denoted by S, is the set/collection of all possible outcomes of that experiment. Usually, we denote an outcome/element of S by ω

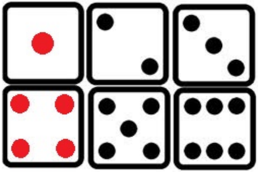
Example 1: Sample Spaces

The *simplest experiment* to which probability (prob) applies is one with *2 possible outcomes*, *e.g.*,

1. When we flip/toss a coin & see which side faces up, $S = \{H, T\}$, where H & T denote head facing up & tail facing up, respectively
2. In an experiment of examining a light bulb to see if it is defective, $S = \{$, $\}$

Example 2: *Sample Spaces*

1. When we roll a die & see the upturned #, the sample space is $S = \{1, 2, \ldots, 6\}$



2. If we examine 3 light bulbs in sequence & note the result of each examination, then an outcome for this experiment is any sequence of 's & 's of length 3, so

$S = \{$  $\}$

3. The amount of time between successive customers arriving at a check–out counter is of interest. For such an experiment,

$$S = \{t : t \geq 0\}$$

Events

- To develop a *mathematical framework* for understanding an experiment & studying prob, *set theory* is naturally adopted in the mathematical theory of prob
- Interested in not only the individual outcomes of S but also various *collections of outcomes* from S

Definition

An *event* A is any collection of outcomes contained in the sample space (*i.e.*, any *subset* of S written as $A \subset S$). An event is *simple* if it consists of exactly 1 outcome & *compound* if it consists of > 1 outcome

- When an experiment is performed & an outcome $\omega \in S$ is *observed/realized*:
  - An event A is said to *occur* if the *observed outcome* ω is included in A (*i.e.*, $\omega \in A$)
  - *Exactly 1 simple event* occurs, but many compound events occur *simultaneously*

Example 3: Events

- **Toss 2 coins**: $S = \{HH, HT, TH, TT\}$
  The event that 2 heads are observed, $A = \{HH\}$, is a *simple event* as ❶ $A \subset S$ & ❷ A contains only 1 outcome
  The event that 1 head is observed, $B = \{HT, TH\}$, is a *compound event*
- **Examine 3 fuses**: *Different compound events* include, *e.g.*,

$A = \{$ , ,  $\}$

$=$ the event that exactly 1 of the 3 fuses is not defective

$B = \{$ , , ,  $\}$

$=$ the event that at most 1 fuse is defective

$C = \{$ ,  $\}$

$=$ the event that all 3 fuses are in the same condition

Suppose that the experiment is performed, & the observed outcome is . Then, the simple event $\{$  $\}$ has occurred & so have the compound events B & C (but not A)

Algebra of Events I

- As an event is just a set, so *set operations* from elementary *set theory* carry over directly into prob theory, which allows us to *create new & more "complex" events from given events*

Definition

Given any event $E, F \subset S$,

1. The *complement* of E, denoted by E^c, $\overline{A}$, or E', is the set of all outcomes in S that are not contained in E
2. The *union* of A & B, denoted by $E \cup F$ & read "E or F", is the event consisting of all outcomes that are *either in A or B or in both events*
3. The *intersection* of E & F, denoted by $E \cap F$ (or, simply EF) & read "E & F", is the event consisting of all outcomes that are *both in E & F*

- *Commutative laws*: $E \cup F = F \cup E$ & $E \cap F = F \cap E$

Algebra of Events II

Definition

Let $\varnothing$ denote the *null event*, the event *consisting of no outcomes whatsoever*

Definition

When $E \cap F = \varnothing$, E & F are said to be *mutually exclusive* or *disjoint* events. It follows that E & E^c must be disjoint for any event $E \subset S$

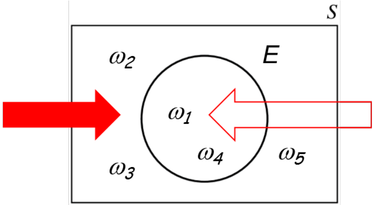
Definition

When $E \cup F = S$, E & F are said to be *exhaustive* events. It follows that E & E^c must be exhaustive

Venn Diagrams I

- *Venn diagrams*: a useful tool for visualizing set operations

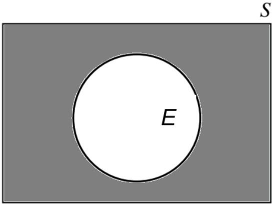
All outcomes $\omega \in S$ are inside the rectangle S



All outcomes $\omega \in E$ are inside the circle representing E

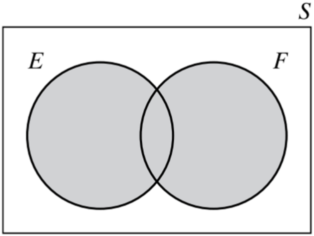
Note: Events can be represented by objects of *any shape*

- *Complement*: E^c is the shaded area

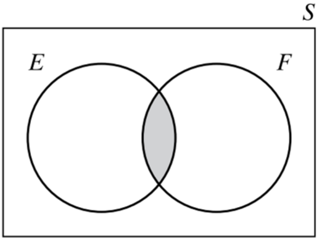


Venn Diagrams II

- *Union*: $E \cup F$ is the <u>shaded area</u>

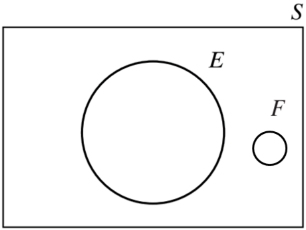


- *Intersection*: $E \cap F$ is the <u>shaded area</u>

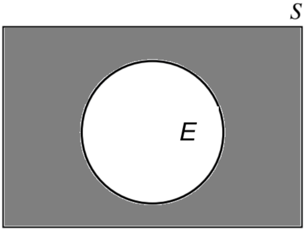


Venn Diagrams III

- E & F are *mutually exclusive or disjoint events*:

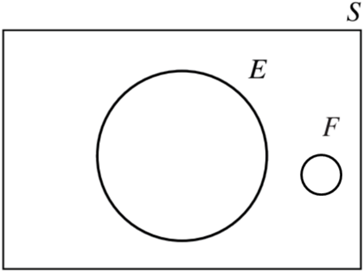


- E & E^c (represented by the shaded area) are *mutually exclusive* & *exhaustive events*:



Example 4: *Venn Diagrams*

- **Toss 2 coins**: $S = \{HH, HT, TH, TT\}$; $F = \{HH\}$ is the right circle; $E = \{HT, TH\}$ is the left circle



How about $C = \{TT\}$?

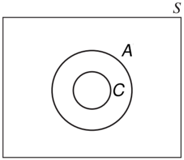
Example 5: *Venn Diagrams & Algebra of Events*

- Assume that there are 7 possible outcomes in an experiment such that (s.t.)

$$S = \{0, 1, 2, 3, 4, 5, 6\}$$

- Let $A = \{0, 1, 2, 3, 4\}$, $B = \{3, 4, 5, 6\}$ & $C = \{1, 2\}$. Then,

$$\begin{aligned}
A^c &= \{5, 6\} \\
A \cup B &= \{0, 1, 2, 3, 4, 5, 6\} = S \text{ (}\textit{i.e.,}\ A \text{ \& } B \text{ are } \textit{exhaustive}\text{)} \\
A \cup C &= \{0, 1, 2, 3, 4\} = A \text{ (indeed, } C \text{ is a } \textit{subset} \text{ of } A\text{)} \\
A \cap B &= \{3, 4\} \\
A \cap C &= \{1, 2\} \\
(A \cap C)^c &= \{0, 3, 4, 5, 6\} \\
B \cap C &= \varnothing \text{ (}\textit{i.e.,}\ B \text{ \& } C \text{ are } \textit{disjoint}\text{)}
\end{aligned}$$



Extensions to ≥ 3 *Events*

- The operations of union & intersection can be extended to ≥ 3 *events*, so can be the idea of disjointness & exhaustiveness
- For any 3 events, $A, B, C \subset S$,
  - $A \cup B \cup C$ is the event consisting of all outcomes that are in at least 1 of the 3 events
  - $A \cap B \cap C$ is the event consisting of all outcomes that are in all the 3 events
  - A, B & C are said to be *mutually exclusive or pairwise disjoint* if no 2 events have any outcomes in common
  - A, B & C are said to be *exhaustive* if the event $A \cup B \cup C$ consists of all outcomes in S

Some Useful Laws

Given any event $A, B, C, E_1, \ldots, E_n \subset S$,

1. *Associative Laws*:

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

2. *Distributive Laws*:

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

3. *DeMorgan's laws*:

$$\left(\bigcup_{i=1}^{n} E_i\right)^c = \bigcap_{i=1}^{n} E_i^c \qquad \left(\bigcap_{i=1}^{n} E_i\right)^c = \bigcup_{i=1}^{n} E_i^c$$

Axioms of Probability

Definition

Probability of an event E is defined by a function P from subsets of S to the real numbers that satisfies the following *axioms of probability*

1. If $E \subset S$, then $P(E) \geq 0$
2. $P(S) = 1$
3. For any countably many sequence of disjoint events $E_1, E_2, \ldots, E_n, \ldots$,
$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$$

- The above axioms *do not* completely determine an assignment of probabilities (probs) to events. They serve only to rule out assignments inconsistent with our *intuitive notions of prob*

Relative–Frequency Interpretation of Probability

Definition of Probability Based on Relative Frequency

Prob of any event E can be defined in terms of its *relative frequency*, *i.e.*, when the same experiment is repeatedly performed for n times under exactly the same conditions, then

$$P(E) = \lim_{n\to\infty} \frac{\text{\# of times that } E \text{ occurs}}{n}$$

- $P(E)$ is defined as the *"limiting"* proportion of time that E occurs, or the "limiting" relative frequency of E
- *Intuitively pleasing* & *theoretically correct*
- Serves as an *interpretation of any prob*

Some More Properties Regarding Probability

Some Simple Propositions Regarding Probabilities

Given any event $A, B \subset S$,

1. $P(\varnothing) = 0$ (*i.e.,* prob that there is no outcome at all is zero)
2. $P(A^c) = 1 - P(A)$
3. $P(A) \leq 1$
4. If $A \subset B$, then $P(A) \leq P(B)$
5. *Addition Law*:

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

- The *addition law* serves as the general formula to compute prob of an event of interest which is expressible as a union of events via decomposing the event of interest into "smaller" events

Example 6: Addition Law

In an undergraduate module, 60% of all students have statistics background, 80% have calculus background, & 50% of all students have both. If a student is selected at random, what is the prob that s/he has background in ❶ at least 1 subject & ❷ exactly 1 subject?

Solution: Let $\begin{cases} A = \{\text{a selected student has statistics background}\} \\ B = \{\text{a selected has calculus background}\} \end{cases}$

We are given: $P(A) = .6$, $P(B) = .8$, & $P(AB) = .5$

P(has background in at least 1 subject)

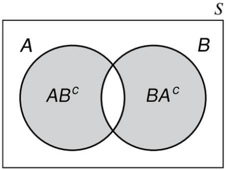
$= P(A \cup B) = P(A) + P(B) - P(AB)$

$= .6 + .8 - .5 = .9$

P(has background in exactly 1 subject)

$= P(AB^c) + P(BA^c) = P(A \cup B) - P(AB)$

$= .9 - .5 = .4$



How to Compute Probabilities?

- Depends on the nature of the sample space S: *finite* versus *infinite*
- When **the sample space is finite** (*i.e.*, $S = \{\omega_1, \omega_2, \ldots, \omega_N\}$):
  - let

$$P(\{\omega_i\}) = p_i, \qquad i = 1, 2, \ldots, N$$

    where $N \geq 1$, called the *cardinality of S*, is a finite positive (+ve) integer denoting the total # of outcomes in S/of the experiment
  - computing $P(A)$ is straightforward due to **Axiom ❸ of prob**:

$$P(A) = P\left(\bigcup_{\omega_i \in A} \{\omega_i\}\right) = \sum_{\omega_i \in A} P(\{\omega_i\}) = \sum_{\{i=1,\ldots,N:\omega_i \in A\}} p_i \tag{1}$$

Simply add probs of all the ω_i that are included in A!

Sample Spaces Having Equally Likely Outcomes

- Assume that
  1. *the sample space is finite* (*e.g.*, $S = \{1, 2, \ldots, N\}$)
  2. *all the N outcomes are equiprobable or equally likely to occur* (a reasonable attribute in many experiments)

Computing Probabilities For Sample Spaces Having Equally Likely Outcomes

For an experiment satisfying the above 2 assumptions, the prob of any event A is

$$P(A) = \frac{\text{\# of outcomes in } A}{\text{\# of outcomes in } S}$$

Simply count the total #'s of outcomes in A & in S!

Example 7: Tossing 2 Fair Coins

- $S = \{HH, HT, TH, TT\}$ with cardinality $N = 4$
- Let A denote the event that at least 1 head is observed
- As $A = \{HH, HT, TH\} = \{HH\} \cup \{HT\} \cup \{TH\}$

$$P(A) = P(\{HH\}) + P(\{HT\}) + P(\{TH\})$$

It remains to find the 3 individual probs

- *Furthermore*, if the coin is *fair* (*i.e.,* equally likely to observe head or tail in any toss), *all the 4 possible outcomes should be equally likely to occur* &

$$P(\{HH\}) = P(\{HT\}) = P(\{TH\}) = P(\{TT\}) = \frac{1}{4}$$

Then,

$$P(A) = \frac{\text{\# of outcomes in } A}{\text{\# of outcomes in } S} = \frac{3}{4}$$

Counting Methods

- In reality, many experiments may be associated with a *finite sample space*, but the *cardinality N could be quite large*
- Understanding S (which involves listing all the outcomes S) is not an easy task; sometimes, even obtaining N, the cardinality of S, can be prohibitive
- Even when the outcomes are equally likely, obtaining the cardinality of A can also be prohibitive
- Several *counting methods or systematic ways for enumeration* will be introduced
  - Especially useful in computing probs when the sample space is finite
  - Some of the ideas can be borrowed or generalized to experiments with *countably infinite* or *uncountably infinite* sample spaces

Basic Principle of Counting

- To deal with "*complex*" experiments defined as *a sequence of several "simpler" experiments*

Basic Principle of Counting & Its Generalization

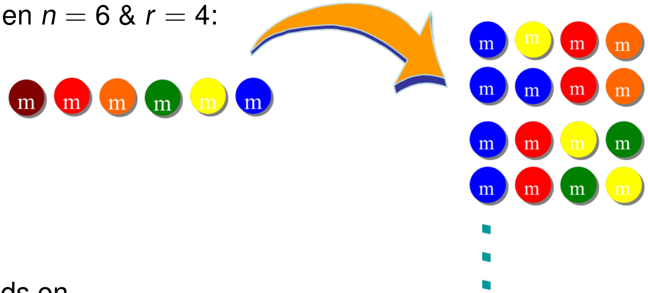
1. If 1 experiment has $m > 0$ outcomes & another experiment has $n > 0$ outcomes, then there are $m \times n$ possible outcomes for the 2 experiments
2. If there are $p > 2$ experiments & the first experiment has n_1 possible outcomes, the second $n_2, \ldots,$ & the pth n_p possible outcomes, then there are a total of $n_1 \times n_2 \times \cdots \times n_p$ possible outcomes for the p experiments

<u>Example 8</u>: *Basic Principle of Counting*

- **<u>Draw a card from a deck of playing cards</u>**: 2 experiments with $m = 4$ & $n = 13$ outcomes (as a card is defined by 4 different suits, ♠, ♥, ♣, ♦, & 13 different face values, A, K, Q, J, 10, 9, ... , 3, 2)
  $\Rrightarrow 4 \times 13 = 52$ possible outcomes

- **<u>Singapore Sweep on first Wednesday of every month</u>**:
  7 experiments with $n_1 = n_2 = \cdots = n_7 = 10$ outcomes
  $\Rrightarrow n_1 \times n_2 \times \cdots \times n_7 = 10^7 = 10,000,000$ possible outcomes
- **<u>Examine m light bulbs</u>**: m experiments with $n_1 = n_2 = \cdots = n_m = 2$ outcomes ( or ) $\Rrightarrow n_1 \times n_2 \times \cdots \times n_m = 2^m$ possible outcomes

- Very often, we would like to address how many ways there are to *select a subset of size r from a group of n* *distinct/distinguishable* *objects* $\{c_1, c_2, \ldots, c_n\}$
- *e.g.,* when $n = 6$ & $r = 4$:



- It depends on
  1. whether we are allowed to *duplicate objects*:
     Sampling without replacement versus *Sampling with replacement*
  2. whether the *sequence/order* from which the r objects are selected matters or not

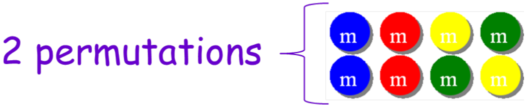
Permutation

A *permutation* is an *ordered arrangement* of objects. Selecting a sample of size r $(= 0, 1, \ldots, n)$ from a set of size n objects, there are

1. n^r permutations under *sampling with replacement*
2. ${}_nP_r \equiv \dfrac{n!}{(n-r)!} = n(n-1)(n-2)\cdots(n-r+1)$ permutations under *sampling without replacement*

- $n!$ is called *n factorial*
- When n is a +ve integer, $n! = n(n-1)(n-2)\cdots 1$. We have the *convention* $0! = 1$
- The *total # of ways in arranging n distinct objects* is $\dfrac{n!}{(n-n)!} = n!$

Example 9: *Permutations*



Refer to the picture at Page 29, how many different ordered arrangements of 4 M&M's selected from 6 M&M's of different colors are there?

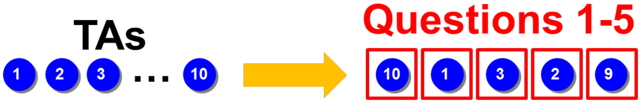
Solution: Here, $n = 6$ *distinct* M&M's/colors (group size) & $r = 4$ selected M&M's/colors (subset size). Note that the order of the 4 colors *matters* & the M&M's cannot be duplicated (*i.e.*, *sampling without replacement*). The # of permutations is

$$\frac{6!}{(6-4)!} = \frac{6!}{2!} = 6(5)(4)(3) = 360$$

Remark: This is equivalent to the *total # of ways to assign 1 out of 6 pieces of distinct M&M's to each of 4 different kids*

Example 10: *Permutations*

There are 10 teaching assistants (TAs) available for grading papers in a test. The test consists of 5 questions, & the Professor wishes to select a different TA to grade each question (at most 1 question per assistant). In how many ways can the TAs be chosen for grading?



Solution: Here, this question is interested in finding the # of permutations with $n = 10$ *distinct* TAs (group size) & $r = 5$ selected TAs (subset size). Note that the order of the 5 selected TAs *matters* (as the 5 questions are *different*) & each TA cannot grade >1 question (*i.e.*, *sampling without replacement*). The # of permutations is

$$\frac{10!}{(10-5)!} = \frac{10!}{5!} = 10(9)(8)(7)(6) = 30,240$$

Example 11: Permutations

There are n rewards to be distributed randomly by a teacher to n students so that each student gets 1 reward. Suppose that 1 of the rewards is the top prize. The students will queue up to receive a reward from the teacher. Shall a student queue first so as to increase his/her chance of getting the top prize?

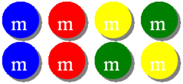
Solution: First of all, all possible assignments of n rewards are equally likely to happen and there are $n!$ # of ways/permutations to distribute the rewards

Assume that a student queues at the i-th ($i = 1, \ldots, n$) position in the queue. Imagine that the n rewards also line up and will be distributed one-by-one accordingly. Now, for this student to have the top prize, the i-th reward should be at the i-th position. There are $(n-1)!$ # of ways to distribute the other $n-1$ rewards into the other $n-1$ positions. Hence, the required prob is

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

Combinations

- Sometimes, we may be *no longer interested in how the objects are arranged or selected*, but in the *constituents of the subset*. For instance, the following permutations are *identical*:



Combination

A *combination* is an *unordered arrangement/collection* of objects. For a set of n *distinct* objects & a subset of size r, there are

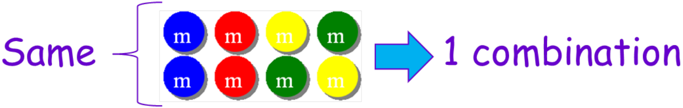
${}_nC_r \equiv \binom{n}{r} \equiv \dfrac{n!}{(n-r)!r!}$ different combinations under *sampling without replacement* when $r = 0, 1, \ldots, n$

- The expression, ${}_nC_r$ or $\binom{n}{r}$, is read as "n choose r" & also called *binomial coefficient*

Example 12: Combinations

Refer to the picture at Page 29, how many different combinations of 4 M&M's selected from 6 M&M's of different colors are there?

Solution:



Here, $n = 6$ *distinct* M&M's/colors (group size) & $r = 4$ selected M&M's/colors (subset size). The order of the 4 colors does *not* matter & we cannot duplicate the M&M's (*i.e.*, *sampling without replacement*). The # of combinations is

$$\frac{6!}{(6-4)!4!} = \frac{6!}{2!4!} = 15$$

which is << 360 (the # of permutations obtained at Page 31)

Example 13: *Bridge Hand*

- A bridge hand consists of any 13 cards selected from a 52–card deck without regard to order, wherein the 52 cards are *distinct*
- There are $\binom{52}{13} = \dfrac{52!}{13!39!} = 635,013,559,600$ different bridge hands
- Since there are 13 cards in each of the 4 suits (♠, ♥, ♣, ♦), the # of hands consisting entirely of ♠ or ♣ (no ♥ or ♦ cards) is $\binom{26}{13} = \dfrac{26!}{13!13!} = 10,400,600$
- Suppose that a bridge hand is dealt from a well–shuffled deck (*i.e.*, every bridge hand/combination of 13 cards from S of cardinality $\approx$ 635 billion is equally likely to occur). So,

$$P(A) = P(\text{obtain a hand consisting entirely of } \spadesuit \text{ or } \clubsuit)$$

$$= \frac{\text{cardinality of } A}{\text{cardinality of } S} = \frac{\binom{26}{13}}{\binom{52}{13}} = .0000164$$

Example 14: *Three of a Kind*

A 5–card poker hand is said to be a "three of a kind" if it consists of 3 cards of the same denomination & 2 other cards of 2 different denominations other than the one for the previous 3 cards. What is the prob that a person is dealt a "three of a kind"?

Solution: Note that all $\binom{52}{5}$ possible poker hands are *equally likely* (*i.e.*, *order* from which the 5 cards are drawn are *irrelevant*)

1. **1 out of 13:** A, K, Q, J, 10, 9, ... , 2 $\boxed{\binom{13}{1}}$ $\boxed{\binom{4}{3}}$ **3 out of 4** ♠, ♥, ♣, ♦ # of combinations to choose the 3 cards of the same denomination, &
2. $\binom{12}{2}\binom{4}{1}\binom{4}{1}$ # of combinations to choose the other 2 cards

Prob of a "three of a kind" is $\dfrac{\binom{13}{1}\binom{4}{3}\binom{12}{2}\binom{4}{1}\binom{4}{1}}{\binom{52}{5}} \approx .02113$

Conditional Probability: A Motivating Example I

- In the process of grade assignment to students in an undergraduate module, the proportion of students getting  is regulated at 10%

  $\Rrightarrow$ The chance of getting  for anyone is .1
- However, in case it was known that *a student has scored full mark in the final examination*

Would the chance of him/her getting *A* still be .1

Conditional Probability: A Motivating Example II

- Of course, it is **intuitive** that the chance of him/her getting  would probably be $> .1$!
- The grade obtained by a student & his/her score in the final examination are *related*
- Once it is known that the student has scored full mark in the final examination (or *equivalently*, the event that the student has scored full mark in the final examination has occurred), it changes our belief in (indeed, it practically increases) the prob of getting  for the student. Indeed, it *changes the experiment*
- Given an experiment with all known characteristics & known probs of outcomes, suppose now it happens that some additional information about the experiment is available. It *may affect the sample space (i.e., the possible outcomes), & the probs associated with each outcome*

Conditional Probability

Definition

Let A & B be 2 events with $P(B) > 0$. The *conditional probability of A given B* is defined to be

$$P(A|B) = \frac{P(AB)}{P(B)} \in [0, 1]$$

- *"Given B"*: an event B has *already occurred*
- *Reduced sample space* $\Leftrightarrow$ *Only outcomes in B, but not in $B^c \subset S$*, are possible to *occur*
  - the sample space for this *"new" experiment* becomes B rather than S & has a *probably smaller cardinality*
  - it is *possible that it is easy to understand A* once B has occurred as B contains fewer outcomes than S
- *Multiplication law* (for computing prob of an intersection in terms of conditional prob): $P(AB) = P(B)P(A|B) = P(A)P(B|A)$ when $P(A), P(B) \neq 0$

Example 15: Conditional Probability I

Suppose that of all individuals buying a certain digital camera, on the spot, 60% also buy a memory card, 40% buy an extra battery, & 30% buy both. Consider randomly selecting a buyer, & let A be the event that a memory card is purchased & B be the event that an extra battery is purchased. Then, we have $P(A) = .6, P(B) = .4, P(AB) = .3$. Given that the selected buyer purchased an extra battery, prob that a memory card was also purchased is

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{.3}{.4} = .75$$

$$P(A^c|B) = \frac{P(A^cB)}{P(B)} = \frac{P(B) - P(AB)}{P(B)} = \frac{.1}{.4} = .25$$

Note: $P(A|B) \neq P(A)$ & $\underline{P(A|B) + P(A^c|B) = 1}$

Example 16: *Conditional Probability II*

A bin contains 25 light bulbs, of which 5 are good & function at least 30 days, 10 are partially defective & will fail in the second day of use, while the rest are totally defective & won't light up at all. Given that a randomly chosen bulb initially lights up, what is the prob that it will still be working after one week?

Solution: Let G be event that the randomly chosen bulb is in good condition, & T be the event that the randomly chosen bulb is totally defective. This implies that T^c represents the event that the selected bulb is either in good condition or partially defective (*i.e.,* the selected bulb lights up initially). The required conditional prob is

$$P(G|T^c) = \frac{P(GT^c)}{P(T^c)} = \frac{P(G)}{P(T^c)} = \frac{5/25}{15/25} = \frac{1}{3}$$

Note: $P(G|T^c) \neq P(G) = 5/25$

Example 17: Conditional Probability And Reduced Sample Space

A fair coin is flipped twice. All outcomes in

$$S = \{HH, HT, TH, TT\}$$

are equally likely. What is the prob that both flips result in heads given that the first flip does?

Solution: Let $A = \{HH\}$ & $B = \{HH, HT\}$. Then, $P(B) = 2/4 = 1/2$, & $P(AB) = P(\{HH\}) = 1/4$. So,

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{1/4}{1/2} = \frac{1}{2}$$

Alternatively, since B has occurred, the *reduced sample space* is $B = \{HH, HT\}$. Then, $AB = A = \{HH\}$, & hence, $P(A|B) = 1/2$

Law of Total Probability I

Definition

A collection of events $B_1, B_2, \ldots, B_n$ is called a *partition* of size n if
❶ $B_iB_j = \varnothing$ for $i \neq j$ & ❷ $\cup_{i=1}^{n} B_i = S$

- $B_1, B_2, \ldots, B_n$ are called *mutually exclusive* & *exhaustive* events
- The simplest partition of size 2 including any event B is $\{B, B^c\}$

Law of Total Probability

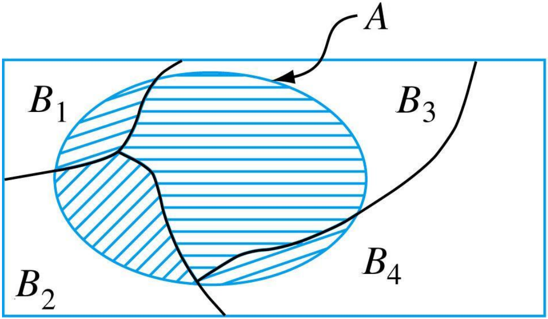
Let $B_1, B_2, \ldots, B_n$ be a partition with $P(B_i) > 0$ for all i. Then, *for any event A*,

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

- To compute any $P(A)$: *Freely choose a partition* s.t. all the *$2n$ probs at the RHS are available or easily computed*

Law of Total Probability II

- *e.g.,* with a partition of size $n = 4$:



- **The idea** :
  1. break A into 4 (or, in general, n) "smaller" events, namely, $A \cap B_1, \ldots, A \cap B_4$ (or $A \cap B_n$)
  2. compute probs of the 4 (or n) intersections by multiplication law

Example 18: *Law of Total Probability*

Of the items produced daily by a factory, 40% come from line 1 & 60% from line 2. Line 1 has a defect rate of 8%, where line 2 has a defect rate of 10%. If an item is chosen at random from the day's production, find the prob that it will not be defective

Solution: Define events: $\begin{cases} D: & \text{item is defective} \\ L_1: & \text{item comes from line 1} \\ L_2: & \text{item comes from line 2} \end{cases}$

From the question, we have $P(L_1) = 1 - P(L_2) = .4$,
$P(D^c|L_1) = 1 - P(D|L_1) = .92$, & $P(D^c|L_2) = 1 - P(D|L_2) = .9$. Then,

$$\begin{aligned} P(\text{an item is not defective}) &= P(D^c) \\ &= P(L_1)P(D^c|L_1) + P(L_2)P(D^c|L_2) \\ &= .4(.92) + .6(.9) \\ &= .908 \end{aligned}$$

Example 19: *Law of Total Probability & Tree Diagram I*

A chain of video stores sells 3 different brands of DVD players. Of its DVD player sales, 50% are brand 1 & 30% are brand 2. Each manufacturer offers 1–year warranty on parts & labor. It is known that 25% of brand 1's players require warranty repair work, whereas the corresponding percentages for brands 2 & 3 are 20% & 10%, respectively. What is the prob that a randomly selected purchaser has a DVD player that will need repair while under warranty?

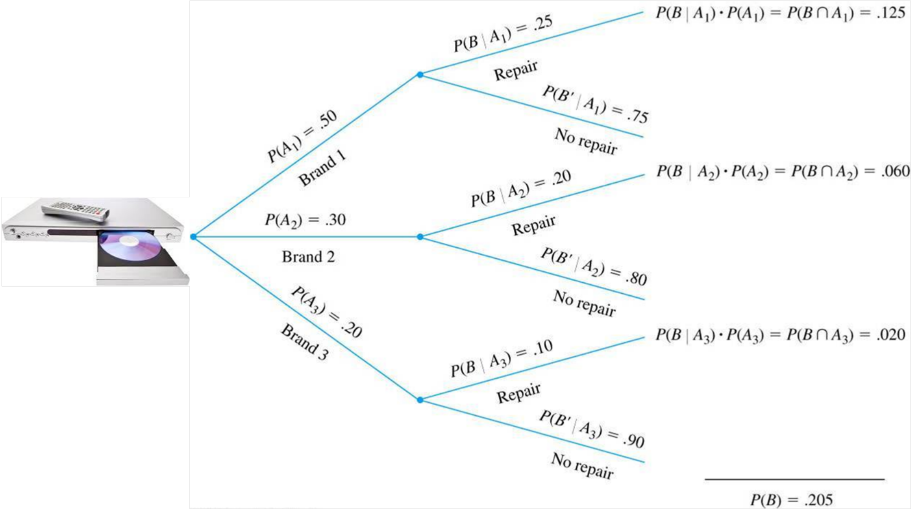
Solution: Let $A_i = \{\text{brand } i \text{ is purchased}\}$, for $i = 1, 2, 3$, & $B = \{\text{needs repair}\}$. Then, we have $P(A_1) = .5, P(A_2) = .3, P(A_3) = .2$, $P(B|A_1) = .25, P(B|A_2) = .2, P(B|A_3) = .1$. Treating *$\{A_1, A_2, A_3\}$ as a partition of size 3* & applying the *law of total prob* yield

$$\begin{aligned} P(B) &= P(\{\text{brand 1 \& repair}\} \text{ or } \{\text{brand 2 \& repair}\} \text{ or } \{\text{brand 3 \& repair}\}) \\ &= P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + P(A_3)P(B|A_3) \\ &= .5(.25) + .3(.2) + .2(.1) = .205 \end{aligned}$$

Example 19: Law of Total Probability & Tree Diagram II

- A *tree diagram* is a handy tool for computing probs in *experiments composing of several stages/generations*
- *Components/ingredients* in a tree diagram:
  - *nodes & branches*; total # in different generations depending on the total # of possible outcomes
  - *probs* attached to each branch
- Here, in this example,
  - the initial/1st generation branches correspond to different brands of DVD players $\Rrightarrow$ 3 branches in the 1st generation
  - the 2nd generation branches correspond to “needs repair” or “doesn’t need repair” $\Rrightarrow$ 2 branches at the end of each branch of the 1st generation
  - probs attached to the branches in the 1st generation
  - probs attached to the branches in the 2nd generation: “conditional probs” based on what has happened in the 1st generation

Example 19: *Law of Total Probability & Tree Diagram III*



Bayes' Formula

Bayes' Formula

Let $B_1, B_2, \ldots, B_n$ be a partition with $P(B_i) > 0$ for all i. Then, for any event A with $P(A) > 0$,

$$P(B_j|A) = \frac{P(B_j)P(A|B_j)}{\sum_{i=1}^{n} P(B_i)P(A|B_i)} \qquad j = 1, 2, \ldots, n$$

- Numerator: $P(AB_j)$; Denominator: $P(A)$
- *Reverse chronological order*: usually B_j happens before A in time
  what should have happened before A has occurred?
- As B_j is the event of interest, it provides a **hint** that one needs to *look for a partition containing B_j* s.t. the *$2n$ probs in the denominator are available or computable*
- **No more complicated than the law of total prob!**

Example 20: Bayes' Formula

Refer to Example 19 about DVD players at Page 47, if a customer returns to the store with a DVD player that needs warranty repair work, what is the prob that it is a brand 1 player? a brand 2 player? a brand 3 player?

Solution: One should apply the Bayes' formula based on the partition $\{A_1, A_2, A_3\}$ to compute these required conditional probs. With $P(B) = .205$ obtained at Page 47, we have

$$P(A_1|B) = \frac{P(A_1)P(B|A_1)}{P(B)} = \frac{.5(.25)}{.205} = .61$$

$$P(A_2|B) = \frac{P(A_2)P(B|A_2)}{P(B)} = \frac{.3(.2)}{.205} = .29$$

$$P(A_3|B) = \frac{P(A_3)P(B|A_3)}{P(B)} = \frac{.2(.1)}{.205} = .1$$

Note: $P(A_1|B) + P(A_2|B) + P(A_3|B) = 1$

Example 21: Calling a Cab

Three telephone lines, A, B & C, are available for calling a cab. The failure rate in connection is 20% for A, 10% for B, & 30% for C. However, line A is more popular & is used for 60% as it is more advertised, whereas line B is used for 30%. A business person failed to connect to a line. What is the prob that s/he used line A?

Solution: Let F be the event of failing to connect to a line. Then,

$$\begin{aligned}
P(A|F) &= \frac{P(FA)}{P(F)} \\
&= \frac{P(A)P(F|A)}{P(A)P(F|A) + P(B)P(F|B) + P(C)P(F|C)} \\
&= \frac{.6(.2)}{.6(.2) + .3(.1) + .1(.3)} \\
&= .667
\end{aligned}$$

Independence I

- An *important concept or notion in prob & statistics* based on conditional probs
- **Recall**: The definition of conditional prob enables us to *revise* the prob $P(A)$ originally assigned to A when we are subsequently informed that another event B has occurred; the "new" prob of A is $P(A|B)$

$$P(A) \stackrel{?}{=} P(A|B)$$

- **Intuitively**, $P(A|B)$ would be *different* from $P(A)$ *unless knowing B does not tell us anything about A* (*i.e.*, "*occurrence of B has nothing to do with occurrence of A*").

Independence II

Definition

Events A & B are said to be *independent (indept)* if

$$
\begin{aligned}
P(AB) &= P(A)P(B), \text{ or equivalently,}\\
P(A) &= P(A|B), \text{ or equivalently,}\\
P(B) &= P(B|A)
\end{aligned}
$$

Otherwise they are said to be *dependent*

- In general, *multiplication law* $P(AB) = P(A)P(B|A) = P(B)P(A|B)$ is *always true for any intersection*; the first formula above, $P(AB) = P(A)P(B)$, is *a special case following from independence (indep) of A & B*
- *Indep* & *disjointness* are *2 very different concepts*
  - conclude disjointness from Venn diagram (no probs involved)
  - indep is defined in terms of probs

Independence III

Indep of 2 Events

If A & B are *indept*, then so are ❶ A & B^c, ❷ A^c & B, & ❸ A^c & B^c

Definition

Three events A, B & C, are said to be *mutually indept* if all the following *4 conditions hold*:

$$
\begin{aligned}
P(ABC) &= P(A)P(B)P(C) \\
P(AB) &= P(A)P(B) \\
P(AC) &= P(A)P(C) \\
P(BC) &= P(B)P(C)
\end{aligned}
$$

- A is indept of any event formed by B & C (*e.g.*, B^c, BC, $B \cup C^c$)
- Three events are *pairwise indept* if the last 3 conditions hold

Example 22: Independence

1. **Draw a card from a deck of playing cards**: Let $A = \{\text{it is an ace}\}$ & $D = \{\text{it is a } \diamondsuit\}$. Intuitively, knowing the card being an ace should give no information about its suit. Verify this with: $P(A) = 4/52 = 1/13$, $P(D) = 13/52 = 1/4$ &

$$P(AD) = P(\{\text{it is } \diamondsuit\text{A}\}) = \frac{1}{52} = \frac{1}{13} \times \frac{1}{4} = P(A)P(D)$$

   $\Rrightarrow$ A & D are *indept*

2. **Toss 2 coins**: Suppose that $A = \{\text{1st coin lands a head}\}$, $B = \{\text{2nd coin lands a head}\}$ & $C = \{\text{exactly 1 head is observed}\}$
   - It is clear that $P(B|A) = P(B)$ as whatever happens to the 1st coin does not influence the 2nd coin $\Rrightarrow$ A & B are *indept*
   - Suppose that the coins are *fair* (*i.e.*, $P(A) = P(B) = .5$). Compute $P(C) = P(\{HT, TH\}) = 2/4 = .5$ & thus

$$P(A|C) = \frac{P(AC)}{P(C)} = \frac{P(\{HT\})}{P(\{HT, TH\})} = \frac{.25}{.5} = .5 = P(A)$$

   $\Rrightarrow$ A & C are *indept*

Example 23: Tossing 2 Dices

Toss 2 fair dice:

1. Let A_6 denote the event that the sum of 2 dice is 6, & B denote the event that the 1st die equals 4

   Then, $\begin{cases} A_6 &= \{\text{⚀⚄, ⚁⚃, ⚂⚂, ⚃⚁, ⚄⚀}\} \\ B &= \{\text{⚃⚀, ⚃⚁, ⚃⚂, ⚃⚃, ⚃⚄, ⚃⚅}\} \end{cases}$

   So, $A_6B = \{\text{⚃⚁}\}$, &

   $$\frac{1}{36} = P(A_6B) \neq P(A_6)P(B) = \frac{5}{36} \times \frac{1}{6}$$

   $\Rightarrow A_6$ & B are *dependent*

2. Let A_7 denote the event that the sum of 2 dice is 7. Then, $P(A_7) = P(\{\text{⚀⚅, ⚅⚀, ⚁⚄, ⚄⚁, ⚂⚃, ⚃⚂}\}) = 6/36 = 1/6$ & $P(A_7B) = P(\{\text{⚃⚂}\}) = 1/36$. Hence,

   $$\frac{1}{36} = P(A_7B) = P(A_7)P(B) = \frac{1}{6} \times \frac{1}{6}$$

   $\Rightarrow A_7$ & B are *indept*