Christopher Lum
lum@uw.edu

Lecture09a
Basic Probability

Lecture is on YouTube

The YouTube video entitled ‘TBD’ that covers this lecture is located at TBD.

Outline

-Experiments,Outcomes,Events
-Probability

Data Representation. Average. Spread

Several examples in this chapter consider the set of data points

$$\begin{array}{l} 89\ 77\ 88\ 91\ 88\ 93\ 99\ 79\ 87\ 84\ 86\ 82\ 88\ 89\ 78 \\ 90\ 91\ 81\ 90\ 83\ 83\ 92\ 87\ 89\ 86\ 89\ 81\ 87\ 84\ 89 \end{array} \tag{1}$$

```
Set1 = {89, 77, 88, 91, 88, 93, 99, 79, 87, 84, 86, 82,
   88, 89, 78, 90, 91, 81, 90, 83, 83, 92, 87, 89, 86, 89, 81, 87, 84, 89};
```

We can sort the list in Mathematica using the ‘Sort’ function

```
Set1Sorted = Sort[Set1]
```

```
{77, 78, 79, 81, 81, 82, 83, 83, 84, 84, 86, 86, 87,
 87, 87, 88, 88, 88, 89, 89, 89, 89, 89, 90, 90, 91, 91, 92, 93, 99}
```

Graphic Representation of Data

We can graphically represent this dataset in several ways

(OPTIONAL) Example 2: Stem-and-Leaf Plot

We first identify N bins to sort the data into. In this example, we choose $N = 5$ and identify the bins as

bins

75-79
80-84
85-89
90-94
95-99

From these bins, we then identify the integer in the 10’s place, this makes up the stems

bins	stems
75-79	7
80-84	8
85-89	8
90-94	9
95-99	9

We now place each individual sample in each stem. For example, the samples 77, 78, and 79 go in the bin 75-70. However, stem 7 already signifies that this bin holds values with 10’s value of 70, we can identify these 3 samples in the bin as 7, 8, 9 or shorthand of 789. These make up the leaves. We can repeat this for each stem to obtain

bins	stems	leaves
75-79	7	789
80-84	8	1123344
85-89	8	6677788899999
90-94	9	001123
95-99	9	9

The number of times that a value occurs is called its **absolute frequency**. Thus 78 has absolute frequency 1, the value 89 has absolute frequency 5, etc. The column to the extreme left in Fig. 507 shows the **cumulative absolute frequencies**, that is, the sum of the absolute frequencies of the values up to the line of the leaf. Thus, the number 10 in the second line on the left shows that our data set has 10 values up to and including 84. The number 23 in the next lines show s that there are 23 values not exceeding 89, etc. Dividing the cumulative absolute frequencies by n ($n = 30$ in this example) gives the **cumulative relative frequencies** 0.1, 0.33, 0.76, 0.93, 1.00.

Leaf unit = 1.0

3	7	789
10	8	1123344
23	8	6677788899999
29	9	001123
30	9	9

Fig. 507. Stem-and-leaf plot of the data in Example 1

Mathematica provides a ‘StemLeafPlot’ function (part of the StatisticalPlots package)

```
Needs["StatisticalPlots`"]
StemLeafPlot[Set1]
```

Stem	Leaves
7	789
8	11233446677788899999
9	0011239

Stem units: 10

As we can see, by default, this creates bins of 70-79, 80-89, and 90-99 which is slightly different than our example bins. Also notice that the ‘Stem units: 10’ means that the stem values of 7, 8, and 9 should be multiplied by 10 to obtain their relative size (or 70, 80, 90, respectively).

We can use several options to customize the Stem-Leaf plot

```
StemLeafPlot[Set1, IncludeStemCounts → True]
```

Stem	Leaves	Counts
7	789	3
8	11233446677788899999	20
9	0011239	7

Stem units: 10

Example 3: Histogram

Another popular method of graphically representing the data set is to use a histogram, which displays the relative class frequency $f_{\text{rel}}(x)$ on the y axis and the bin on the x axis

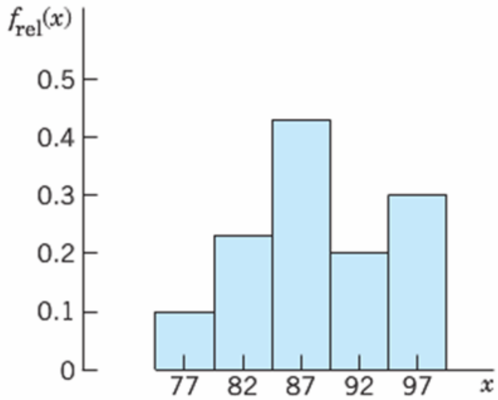
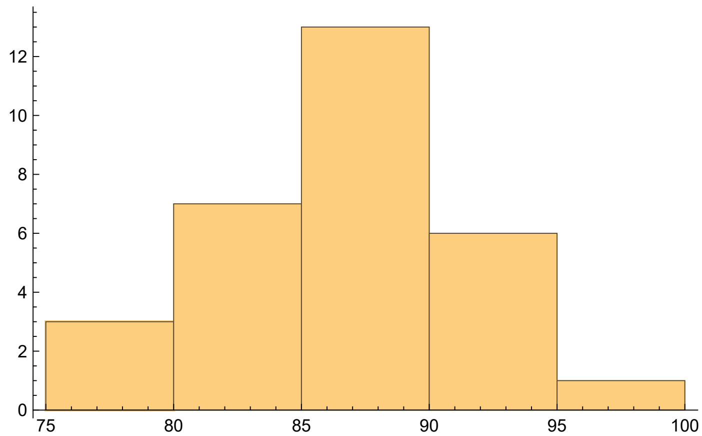


Fig. 508. Histogram of the data in Example 1 (grouped as in Fig. 507)

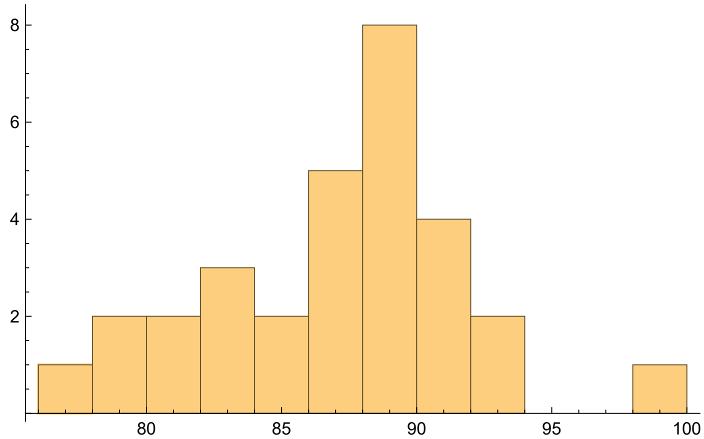
Mathematica provides the 'Histogram' function to plot a histogram

```
Histogram[Set1]
```



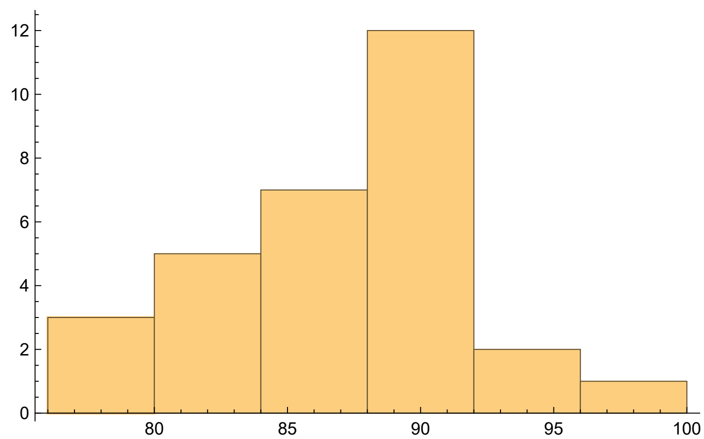
We can set options on how we want to bin the histogram. For example, we can use 10 bins.

```
numBins = 10;
Histogram[Set1, numBins]
Clear[numBins]
```



Or we can specify bin width

```
binWidth = 4;
Histogram[Set1, {binWidth}]
Clear[binWidth]
```



Matlab provides the 'hist' function to perform similar operations with histograms.

(OPTIONAL) Example 4: Boxplot. Media. Interquartile Range. Outlier

A **boxplot** of a set of data illustrates the average size and spread of the values.

The average size is measured by the **median**, or middle quartile, q_M.

$$q_M = \begin{cases} \text{middlemost values of the sorted data set} & \text{if } n \text{ is odd} \\ \text{average of the two middlemost values of the sorted data set} & n \text{ is even} \end{cases}$$

In our situation, we have an even number of samples, so the median is

$$q_M = \frac{1}{2}(x_{15} + x_{16}) = \frac{1}{2} \times (87 + 88) + 87.5$$

Mathematica provides the 'Median' function to compute this

```
Median[Set1] // N
```

```
87.5
```

The spread of values is measured by the **range**, *R*, which is the difference between the largest and smallest sample

$$R = x_{\max} - x_{\min}$$

```
Max[Set1] - Min[Set1]
```

```
22
```

Better information about the spread of data can be obtained from the **interquartile range**, IQR

$$\text{IQR} = q_U - q_L$$

where q_U =median of data above the median (q_M)
q_L = median of data below the median (q_M)

In our example, we can compute $q_L = x_8 = 83$ and $q_U = x_{23} = 89$

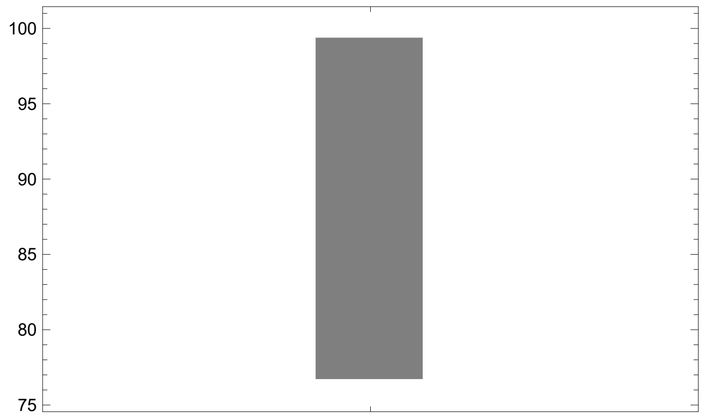
```
(*Break the sorted list into an upper and lower set*)
If[OddQ[Length[Set1Sorted]],
 (*----------------Odd number of samples--------------*)
 (*Identify the middle element*)
 middleIndex = Floor[Length[Set1Sorted] / 2] + 1;

 (*Remove this from the list*)
 Set1SortedEven = Delete[Set1Sorted, middleIndex];

 (*Break the sorted list into an upper and lower set*)
 temp = Partition[Set1SortedEven, Length[Set2Even] / 2],

 (*-------------Even number of samples----------------*)
 temp = Partition[Set1Sorted, Length[Set1Sorted] / 2];
]

lowerSet = temp[[1]];
upperSet = temp[[2]];
qL = Median[lowerSet]
qU = Median[upperSet]
```

```
83
```

```
89
```

Mathematica provides the 'BoxWhiskerChart' function to plot this. Note: This was called 'BoxWhiske-rPlot' in versions lower than 8.0. Also note that you can hover over the chart to see information about the data.

```
BoxWhiskerChart[Set1]
```



Mean. Standard Deviation. Variance. Empirical Rule.

We can compute the average, or **mean** as (note that the bar notation means average, not a vector)

$$\mu = \overline{x} = \frac{1}{n} \sum_{j=1}^{n} x_j \qquad \textbf{\textit{(Eq.3)}}$$

Mathematica provides the 'Mean' function to compute the mean of a list

```
Mean[Set1] // N
```

```
86.6667
```

We can measure the spread of the data by the **standard deviation**, σ. In some applications, it becomes useful to look instead at the standard deviation squared, σ^2, which is also called the **variance**.

$$\sigma^2 = \frac{1}{n-1} \sum_{j=1}^{n} \left(x_j - \overline{x}\right)^2 \qquad \textbf{\textit{(Eq.4)}}$$

$$\sigma^2 = \frac{1}{n} \sum_{j=1}^{n} \left(x_j - \overline{x}\right)^2 \qquad \text{(alternate definition of variance for larger data sets)}$$

Mathematica provides the 'StandardDeviation' and 'Variance' functions to compute this. Note that Mathematica uses Eq.4 (ie divides by n-1) to compute the standard deviation. Other software tools may divide by *n* so be sure to be careful if this matters in your application.

```
StandardDeviation[Set1] // N
Variance[Set1] // N
```

```
4.80182
```

```
23.0575
```

Empirical Rule. For any mound-shaped, nearly symmetric distribution of the data, the intervals $\overline{x} \pm s$, $\overline{x} \pm 2\,s$, $\overline{x} \pm 3\,s$ contain about 68%, 95%, and 99.7% of the points, respectively.

Experiments, Outcomes, Events

We define some terms related to probability theory.

Experiment: A process of measurement or observation. This could be in a laboratory, a factory, etc.

Trial: This is a single performance of an experiment.

Sample Point: This is the result of a trial. This is also referred to as an outcome.

Sample Space: This is the set of all possible outcomes of an experiment. We often denote this set as S

Example 7: Events

Consider the experiment of rolling a fair, 6-sided die.

Experiment: Rolling a fair, 6-sided die

Trial: A single roll of the die

Sample Space: $S = \{1, 2, 3, 4, 5, 6\}$ (all possible outcomes in the experiment)

We can now define some events that can occur in the experiment. For example, we can define an event A as rolling an odd number and B as rolling an even number

$A = \{1, 3, 5\}$ (odd number)
$B = \{2, 4, 6\}$ (even number)

We can define an event C as rolling a 1 or 2

$C = \{1, 2\}$ (roll a 1 or 2)

A simple event is one such as $\{1\}$, $\{2\}$, ..., $\{6\}$ that only has a single value.

In a trial, if outcome a happens and $a \in A$ (a is an element of event/set A), we say that event A happens.

For example, if we perform a single trial and the result is the die comes up as a 2, then both events B and C occurred.

Unions, Intersections, Complements of Events

Since events are basically sets, much of the same set theory can be applied to events.

The **union,** $A \cup B$, of A and B consists of all points that are in either A or B

The **intersection**, $A \cap B$, of A and B consists of all points that are in both A and B.

If A and B have no points in common, we write $A \cap B = \emptyset$ (empty set) and we call A and B **mutually**

exclusive, because A happening means that B cannot happen as well, and vice versa. For example, in our previous example, A and B are mutually exclusive because if you roll an odd number, you cannot simultaneously roll an even number, and vice versa.

The **complement**, A^c, is the set of all points of S that are not in A. Therefore,

$$A \cap A^c = \emptyset$$
$$A \cup A^c = S$$

It may be helpful to visualize sets using **Venn diagrams**. This is a diagram that shows all possible logical relations between a finite collection of different sets.

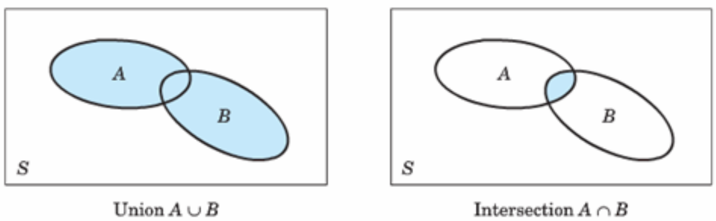


Fig. 510. Venn diagrams showing two events A and B in a sample space S and their union $A \cup B$ (colored) and intersection $A \cap B$ (colored)

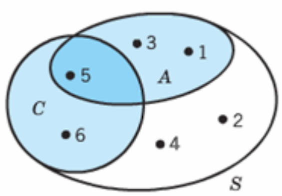


Fig. 511. Venn diagram for the experiment of rolling a die, showing S, $A = \{1, 3, 5\}$, $C = \{5, 6\}$, $A \cup C = \{1, 3, 5, 6\}$, $A \cap C = \{5\}$

Example 8: Unions and Intersections of 3 Events

Consider our experiment of rolling a die. We define 3 events

$$A = \text{number greater than } 3 = \{4, 5, 6\}$$
$$B = \text{number less than } 6 = \{1, 2, 3, 4, 5\}$$
$$C = \text{even number} = \{2, 4, 6\}$$

```
A = {4, 5, 6};
B = {1, 2, 3, 4, 5};
Cc = {2, 4, 6};  (*need to use Cc because C is a reserved symbol*)
```

We can now compute various intersections and unions using Mathematica. Note that there are multiple ways to do this in Mathematica. Note that the shortcut for ∪ and ∩ is \bigcup and \bigcap, respectively.

```
Print["A⋃B"]
A⋃B
Print[]

Print["Union[A,B]"]
Union[A, B]
Print[]

Print["A⋂B"]
A⋂B
Print[]

Print["Intersection[A,B]"]
Intersection[A, B]
Print[]
```

```
A⋃B

{1, 2, 3, 4, 5, 6}


Union[A,B]

{1, 2, 3, 4, 5, 6}


A⋂B

{4, 5}


Intersection[A,B]

{4, 5}
```

Probability

Definition 1: First Definition of Probability

If the sample space S of an experiment consists of finitely many outcomes (points) that are equally likely, then the probability $P(A)$ of an event A is

$$P(A) = \frac{\text{number of points in } A}{\text{number of points in } S} \qquad \textbf{\textit{(Eq.1)}}$$

From this definition, it follows immediately that

$$P(S) = 1 \qquad \textbf{\textit{(Eq.2)}}$$

In many cases, the number of points in S or A may be too large to enumerate (S = flavors of ice cream and A = brown ice creams). In these cases, it may be easier to experimentally relate the probability to the relative frequency of an event by defining the **relative frequency** of A in a set of trials as

$$f_{\text{rel}}(A) = \frac{f(A)}{n} = \frac{\text{number of times } A \text{ occurs}}{\text{number of trials}} \qquad \textbf{\textit{(Eq.3)}}$$

where $\quad f(A)$ = absolute frequency of event A

If A and B are mutually exclusive, they cannot occur together. Hence the absolute frequency of their union $A \cup B$ must equal the sum of the absolute frequencies of A and B.

$$f_{\text{rel}}(A \cup B) = \frac{f(A \cup B)}{n}$$ note: because A and B are mutually exclusive, they cannot occur together

$$= \frac{f(A)+f(B)}{n}$$

$$= \frac{f(A)}{n} + \frac{f(B)}{n}$$

$$f_{\text{rel}}(A \cup B) = f_{\text{rel}}(A) + f_{\text{req}}(B) \qquad \text{if } A \cap B = \varnothing \qquad \textbf{\textit{(Eq.6*)}}$$

For example, if event A is rolling a 1 on a die and event B is rolling a 2 or 3, we see that $f_{\text{rel}}(A \cup B) = f_{\text{rel}}(1 \text{ or } 2 \text{ or } 3) = 3/6$ and we see that by themselves, the relative frequency of each event is $f_{\text{rel}}(A) = 1/6$ and $f_{\text{rel}}(B) = 2/6$.

Definition 2: General Definition of Probability

Given a sample space S, with each event A of S (subset of S) there is associated a number $P(A)$ called the **probability** of A, such that the following **axioms of probability** are satisfied.

1. For every A in S

$$0 \leq P(A) \leq 1 \qquad \textbf{\textit{(Eq.4)}}$$

2. The entire sample space S has the probability

$$P(S) = 1 \qquad \textbf{\textit{(Eq.5)}}$$

3. For mutually exclusive events A and B (where $A \cap B = \varnothing$)

$$P(A \cup B) = P(A) + P(B) \qquad \text{for } A \cap B = \varnothing \qquad \textbf{\textit{(Eq.6)}}$$

Basic Theorems of Probability

We can now form some basic theorems of probability.

Theorem 1: Complementation Rule

For an event A and its complement A^c in a sample space S

$P(A^c) = 1 - P(A)$ ***(Eq.7)***

Theorem 2: Addition Rule for Mutually Exclusive Events

For mutually exclusive events $A_1, A_2, \ldots, A_m$ in a sample space S

$P(A_1 \cup A_2 \cup \ldots \cup A_m) = P(A_1) + P(A_2) + \ldots + P(A_m)$ ***(Eq.8)***

Theorem 3: Addition Rule for Arbitrary Events

For events A and B in a sample space (not necessarily mutually exclusive)

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$ ***(Eq.9)***

Example 4: Union of Arbitrary Events

In tossing a fair die, what is the probability of getting an odd number or a number less than 4?

We define 2 events

$A = \{1, 3, 5\}$ (odd number)
$B = \{1, 2, 3\}$ (less than 4)

We seek to compute $P(A \cup B)$. We can do this by brute force by first computing $A \cup B$ (the probability of getting and odd number <u>or</u> a number less than 4).

$A \cup B = \{1, 2, 3, 5\}$

Because the die are fair, each outcome is equally likely and we can therefore use Eq.1 to compute the probability

$$P(A \cup B) = \frac{\text{number of points in } A \cup B}{\text{number of points in } S} = \frac{4}{6} = \frac{2}{3}$$

Alternatively, we can use Theorem 3.

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$ note: $P(A) = \frac{3}{6}$, $P(B) = \frac{3}{6}$, and $A \cap B = \{1, 3\} \Rightarrow P(A \cap B) = \frac{2}{6}$

$$= \frac{3}{6} + \frac{3}{6} - \frac{2}{6}$$

$$P(A \cup B) = \frac{2}{3}$$

which is the same solution.

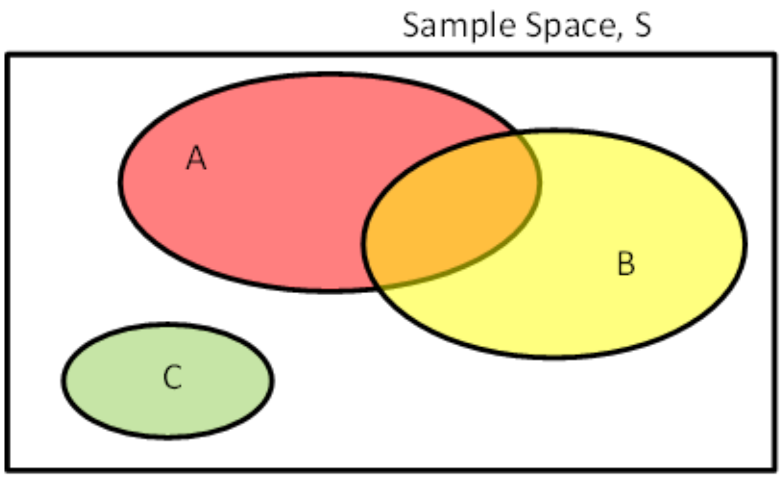
Conditional Probability. Independent Events

Often it is required to find the probability of an event B under the condition that an event A occurs. This probability is called the **conditional probability** of B given A and is denoted by $P(B \mid A)$. In this case, A serves as a new (reduced) sample space, and the conditional probability is the fraction of $P(A)$ which corresponds to $A \cap B$.

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} \qquad \textbf{\textit{(Eq.11)}}$$

where $P(A) \neq 0$

We can illustrate Eq.11 with a Venn diagram



Recall that we seek to obtain $P(B \mid A)$. From the diagram, we see that if we know that A has occurred, then the probability that B has occurred simultaneously is the area of the orange/red intersection divided by the area of the red. In other words $P(B \mid A) = \frac{P(A \cap B)}{P(A)}$ which is what Eq.11 states.

Similarly the conditional probability of A given B is

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \qquad \textbf{\textit{(Eq.12)}}$$

where $P(B) \neq 0$

We examine an example where the solution changes depending on how you pose the problem.

Example: Boys and Girls in a Family
Suppose that a family has 2 children. Assuming there is an equal likelihood of having a boy or girl, there are several questions we would like to answer.

Question 1. What is the probability that the first child is a girl?
Question 2. What is the probability that both children are girls?
Question 3. You learn that at least one of the children is a girl. What is the probability that the other child is a girl as well?

Question 1 is trivial, we see that the probability of the first child being a girl is 1/2.

For question 2, we can enumerate the sample space

$S = \{BB, BG, GB, GG\}$

where BG means that the first born is a boy and the second is a girl.

Assuming that boys are girls are equally likely to be born, the 4 elements of *S* are equally likely. Therefore, we see that

$P(GG) = 1/4$

Question 3 is more nuanced and requires closer inspection.

We can define an event, *E*, that the family has at least 1 girl as the set

$E = \{GG, BG, GB\}$ (family has at least 1 girl)

We define the event *F* that the family has two girls as

$F = \{GG\}$ (family has two girls)

We now want to compute the probability of *F* given that *E* has occurred. In other words, we seek $P(F \mid E)$. We can use Eq.11 to compute this conditional probability.

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

$$= \frac{P(\{GG\})}{P(\{GG, BG, GB\})}$$

$$= \frac{1/4}{3/4}$$

$P(F \mid E) = 1/3$ (probability that the family has two girls given that one is a girl)

Note that this type of problem is affected by the type of assumptions we make during solving it.

Show Matlab code for performing a Monte Carlo style simulation to support this analysis

Side Note: The Monty Hall Problem has some relation to conditional probability (gameshow style setting where contestant is given the option to switch doors after being shown on of the doors that does not contain the prize). ***Perform demonstration of the game***

Solving Eq.11 and Eq.12 for $P(A \cap B)$ gives Theorem 4 as stated below.

Theorem 4: Multiplication Rule

If *A* and *B* are events in a sample space *S* and $P(A) \neq 0$, $P(B) \neq 0$, then

$$P(A \cap B) = P(A)\, P(B \mid A) = P(B)\, P(A \mid B) \qquad \textbf{(Eq.13)}$$

Example 5: Multiplication Rule

In producing screws, let us define two events as

A: "screw too slim"
B: "screw too short"

Let $P(A) = 0.1$ (probability that the screw is too slim) and let the conditional probability that a slim screw is also too short be $P(B \mid A) = 0.2$. What is the probability that a screw that we pick randomly from the lot will be both too slim and too short?

We seek $P(A \cap B)$. Therefore, by Eq.13, we have

$$P(A \cap B) = P(A)\,P(B \mid A) = (0.1) \times (0.2) = 0.02 = 2 \text{ percent}$$

Independent Events

If events A and B are such that

$$P(A \cap B) = P(A)\,P(B) \qquad \textbf{\textit{(Eq.14)}}$$

then A and B are called **independent events**. Assuming that $P(A) \neq 0$, $P(B) \neq 0$, from Eq.13, for independent events we can write

$$P(A \cap B) = P(A)\,P(B \mid A) = P(B)\,P(A \mid B) \qquad \text{recall: } P(A \cap B) = P(A)\,P(B)$$

$$P(A)\,P(B) = P(A)\,P(B \mid A) = P(B)\,P(A \mid B)$$

which leads to

$$P(A \mid B) = P(A) \text{ and } \qquad P(B \mid A) = P(B)$$

This means that the probability of A does not depend on the occurrence or non-occurrence of B (similarly for B).

Note that often people mistakenly interchange mutually exclusive events and independent events. In fact, these two are nearly opposite of one another. If two events is mutually exclusive, then the occurrence of one gives you 100% information about the other event, namely it cannot occur.

$$A \cap B = \varnothing \text{ (definition of mutually exclusive events)} \Rightarrow P(A \cap B) = 0$$

$$P(A \cap B) = P(A)\,P(B) \text{ (definition of independent events)}$$

Independence of *m* Events

Similarly, m events $A_1, A_2, \ldots, A_m$ are called independent if

$$P(A_1 \cap A_2 \cap \ldots \cap A_m) = P(A_1)\,P(A_2) \ldots P(A_m)$$

Accordingly, three events, A, B, C are independent if and only if

$$
\begin{aligned}
&P(A \cap B) = P(A)\,P(B) \\
&P(B \cap C) = P(B)\,P(C) \qquad \textbf{\textit{(Eq.16)}} \\
&P(C \cap A) = P(C)\,P(A) \\
&P(A \cap B \cap C) = P(A)\,P(B)\,P(C)
\end{aligned}
$$

Example: Revisit the Family of Girls Example

Let us revisit the previous example of the family with 2 children. Let us reconsider a question 2 in the framework of independent events

Question 2: What is the probability that the family has two girls?

We can define two events

A = first child is a girl
B = second child is a girl

We easily compute that

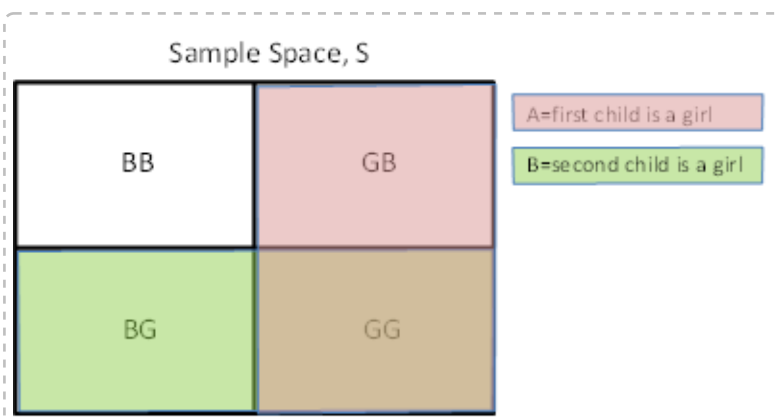
$P(A) = 1/2$
$P(B) = 1/2$

Because the probability of a second child being a girl is independent of the first child being a girl, we simply compute the probability of both the first and second children being girls via Eq.14

$P(A \cap B) = P(A)\,P(B) = \frac{1}{2} * \frac{1}{2} = \frac{1}{4}$ (use multiplication rule because A and B are independent events)

This agrees with the probably you would obtain by simply enumerating the sample space and computing the probability of the {GG} event as we did previously.

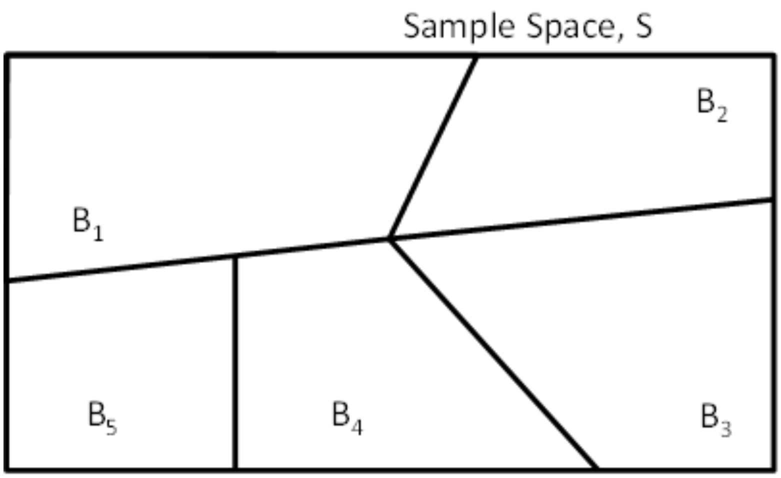
Furthermore, if we draw a Venn diagram we obtain



We see that indeed, independent events must have non-empty intersection and the area rule of $P(A \cap B) = P(A)\,P(B)$ is true.

Law of Total Probability

Consider the situation where we divide the sample space into a finite or countably infinite partition, denoted by $\{B_n : n = 1, 2, 3, \ldots\}$. In other words, these are sets of disjoint events whose union is the entire sample space)



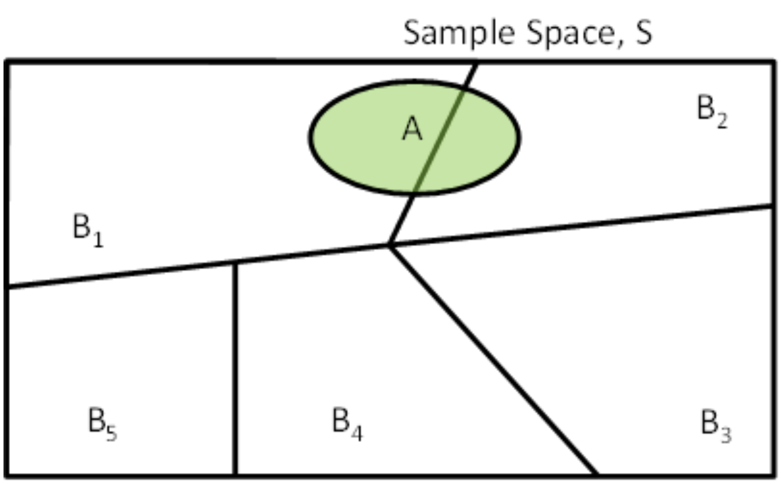
Then for any event, A, in the same sample space

$$P(A) = \sum_n P(A \cap B_n) \qquad \textbf{(Eq.A)}$$

Or alternatively (using the definition of conditional probability) $P(A \cap B_n) = P(A \mid B_n) P(B_n)$

$$P(A) = \sum_n P(A \mid B_n) P(B_n) \qquad \textbf{(Eq.B)}$$

Eq.A is relatively easy to derive/visualize if we draw a Venn diagram



So in this case, we see that $P(A)$ is the area of the green oval. This is obviously equal to

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + P(A \cap B_3) + P(A \cap B_4) + P(A \cap B_5)$$

In this particular example, we see that $P(A \cap B_3) = P(A \cap B_4) = P(A \cap B_5) = 0$

Example

Consider an elementary school with 250 students. The distribution of students are as follows

K : 45
1 : 55
2 : 60
3 : 27
4 : 63

Grade	#Students	#M	#F	#Buying	#NotBuying	comment
K	45	25	20	8	37	□
1	55	33	22	33	22	□
2	60	28	32	28	32	□
3	27	16	11	9	18	□
4	63	35	28	0	63	field trip, no lunch

$P(A \cap H) = \frac{8}{250}$ $P(A \cap I) = \frac{37}{250}$

$P(B \cap H) = \frac{33}{250}$ $P(B \cap I) = \frac{22}{250}$

$P(C \cap H) = \frac{28}{250}$ $P(C \cap I) = \frac{32}{250}$

$P(D \cap H) = \frac{9}{250}$ $P(D \cap I) = \frac{18}{250}$

$P(E \cap H) = \frac{0}{250}$ $P(E \cap I) = \frac{63}{250}$

Consider the following events:

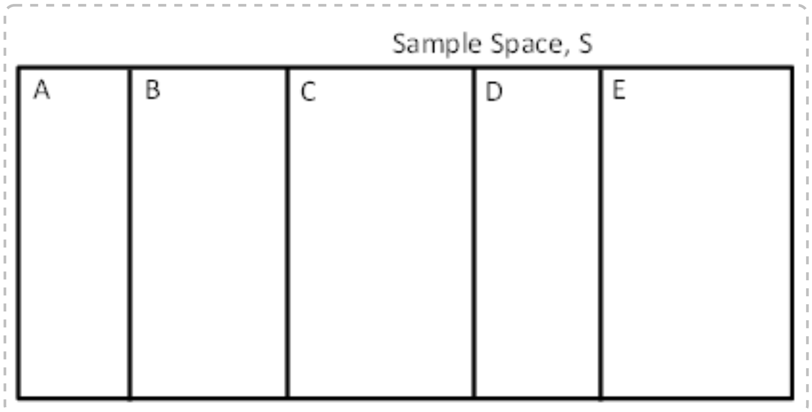
A : the student is in kindergarten
B : the student is in 1st grade
C : the student is in 2nd grade
D : the student is in 3rd grade
E : the student is in 4th grade

F : the student is a boy
G : the student is a girl

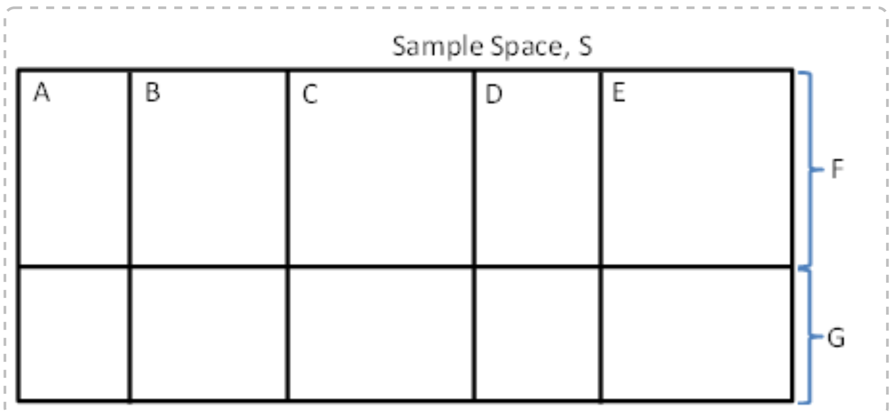
H : the student is buying lunch
I : the student is not buying lunch

Show there are no 4th grade students buying lunch

Part a. Draw a Venn diagram showing events $A - E$



Part b. Update the Venn diagram showing events F and G



Part c. Suppose that the probability of lunch being ordered by any given student is 0.312. Furthermore, we know the probability of K, 1, 2, 3 grades buying lunch are $\frac{8}{250}$, $\frac{33}{250}$, $\frac{28}{250}$, $\frac{8}{250}$, respectively.

Show there are no 4th graders buying lunch.

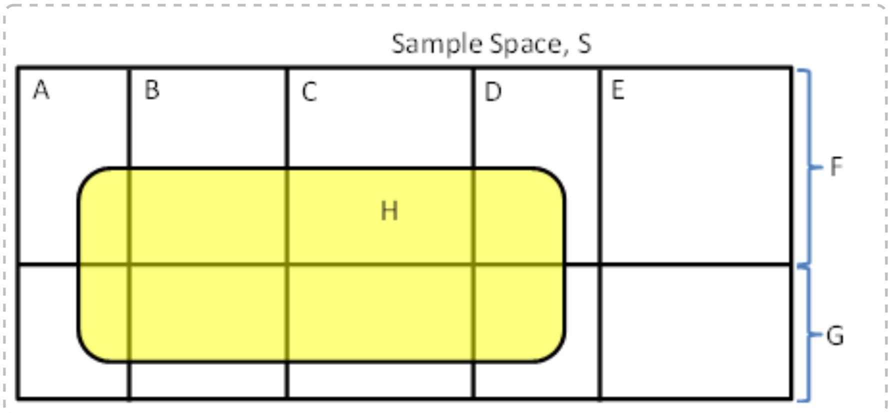
From part a, we know that $\{A, B, C, D, E\}$ forms a partition of the sample space and therefore, by the law of total probability

$$P(H) = P(H \cap A) + P(H \cap B) + P(H \cap C) + P(H \cap D) + P(H \cap E)$$

$$0.312 = \frac{8}{250} + \frac{33}{250} + \frac{28}{250} + \frac{9}{250} + P(H \cap E)$$

$$P(H \cap E) = 0$$

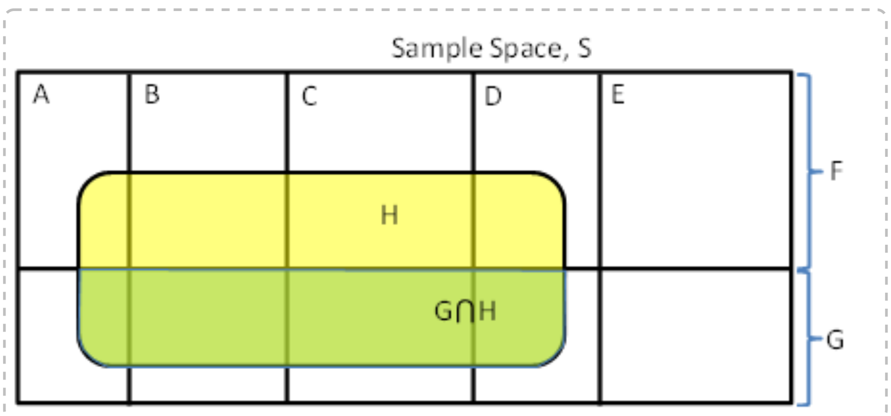
Part d. Update your Venn diagram to show events H and I



Event I is simply the areas that are not yellow.

Part e. Assume the probability of a boy at school buying lunch is 39/250. What is the probability that a girl is buying lunch?

We seek $P(G \cap H)$ which is shown in green below.



We seek $P(H \cap G)$ (buying lunch and is a girl). Because $\{F, G\}$ forms a partition of the space, from the law of total probability we can write

$$P(H) = P(H \cap F) + P(H \cap G)$$

$$P(H \cap G) = P(H) - P(H \cap F)$$

We know that $P(H) = 0.312$ and $P(H \cap F) = \frac{39}{250}$

$$P(H \cap G) = 0.312 - \frac{39}{250}$$

$$P(H \cap G) = 0.156$$