

STA107

An Introduction to Probability

Lecture 1: *Fundamentals*

Winter 2001

Wednesday 7pm-10pm

Ian Buckley*

* With thanks to Prof. David Brenner for use of his course structure, notes and problem sheets.

Some of the problems about Chance having a great appearance of Simplicity, the Mind is easily drawn into a belief, that their solution may be attained by the mere Strength of natural good Sense; which generally proving otherwise and the Mistakes occasioned thereby being not unfrequent, 'tis presumed that a Book of this Kind, which teaches to distinguish Truth from what seems so nearly to resemble it, will be looked upon as a help to good reasoning.

De Moivre in his dedication of *The Doctrine of Chance*

God does not play dice with the universe

Albert Einstein - in disagreement with the Copenhagen school of quantum mechanics, which holds that the universe is manifestly random in nature

1.1 Introduction

1.1.1 Motivation

... probability as a measurable degree of certainty; necessity and chance; moral versus mathematical expectation; a priori and a posteriori probability; expectation of winning when players are divided according to dexterity; regard of all available arguments, their valuation, and their calculable evaluation; law of large numbers ...

Jacob Bernoulli 1654-1705

- "Much of our life based on belief that future is largely unpredictable." (Grimmett and Stirzaker)
- Belief in chance behaviour → use of words *random* or *probability*
- *Quantitative* as well as *qualitative* meaning
- Mathematical theory \iff common rational understanding
- What is probability theory? A branch of mathematics concerned with the analysis of random phenomena. A random event may or may not occur, and whether it does cannot be determined in advance.
- *Independence* between two events - probability of B is not changed by knowing that A has occurred (and vice versa)
- *Conditional probability* - probability of B occurring might be different if we know that A has occurred (if A and B are not independent).
- *Law of large numbers*: experiment repeated many times under identical conditions, outcomes recorded.

The LLN implies that *the proportion of experiments which give rise to some specified outcome is approximately equal to the probability of that outcome*. Provides a means to estimate probabilities by observing relative frequencies of the outcomes in a set of repetitions of an experiment.

- *The central limit theorem* gives information about how the observed relative frequency will differ from the probability of the outcome. Universal probability law: *normal* or *Gaussian* curve.

■ Load *Mathematica* Packages

1.1.2 Applications

■ 1.1.2.1 Gambling

...we are compelled to gamble...

Blaise Pascal 1623-1662

- Original motivation of the subject
- Books
 - *Taking Chances* John Haigh
 - *Can You Win? the Real Odds for Casino Gambling, Sports Betting, and Lotteries*, Mike Orkin
 - *Efficiency of Racetrack Betting Markets*, Donald B. Hausch (Editor), Victor S.Y. Lo (Editor), William T. Ziemba (Editor)

- Nice problem: Gambler's Ruin. Two players, toss a coin. For each head, player 1 pays player 2 \$1; and for each tail, player 2 pays player 1 \$1. If the players have endowments of \$ x and \$ y what is the probability that either of the players will be ruined? Also, how long does the game last?

■ 1.1.2.2 Statistics

- Medical
- Biological
- *Legal Statistics for Lawyers* Michael O Finkelstein, Bruce Levin

■ 1.1.2.3 Pure science

- Physics - statistical mechanics, quantum mechanics
- Nuclear physics and nuclear engineering
- Genetics - the sexes of babies follow sequences similar to those of coin tosses

■ 1.1.2.4 Engineering

- Stochastic control (chemical, aeronautical (Apollo))
- Reliability - Abraham Wald during WW2 - Generalizations of the problem of gambler's ruin; quality control
- Computer science and electrical engineering - information theory, error checking,

■ 1.1.2.5 Business & Government

- Finance
 - insurance - insurance company places a series of bets on the health and life expectancy of people or the integrity of their belongings. Using historical records, the insurer employs probability theory to win more often than he or she loses.
 - risk-neutral pricing of derivatives
 - weather derivatives
 - portfolio management ("mean-variance" approach of Markowitz)
 - risk management, capital adequacy and "Value at Risk"
- Optimal decision making under uncertainty
 - which project to develop
 - how to value flexibility
- Cryptography

■ 1.1.2.6 General numeracy and life

- *Once Upon a Number* John Allen Paulos - statistics of stereotyping and race relations. Draws parallels between statistics and story telling.
- Cost benefit analysis, health scares and hysteria
- Coincidences and debunking superstitions

■ 1.1.2.7 Philosophy

Although the debate still simmers, mathematicians - like losing generals everywhere - have simultaneously retreated and declared victory... Such a definition may not be philosophically gratifying, but is mathematically liberating.

John Allen Paulos Discussing the axiomatisation of probability P68, *Once Upon a Number*

1.1.3 About these notes

1.2. History

1.3. The Basics

1.3.1 Picking Objects With Replacement

Example 1.3.2 Consider a box containing n tickets, numbered $1, 2, \dots, n$.

- Draw a ticket at random ("random" \Rightarrow each has the same chance $= \frac{1}{n}$; the *uniform* distribution)

```
BarChart[Table[{1/5, x}, {x, 0, 5}], AxesLabel -> {"x", "P[X = \ x]"}];
```

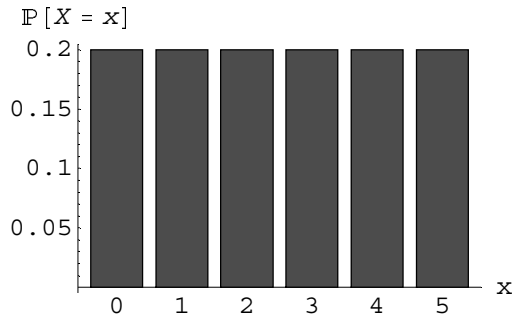


Figure 1

Bar Chart of the Uniform Distribution for number of tickets, $n = 5$.

- Probability that the ticket is, say 1 is $1/n$
- *Replace* into the box
- Draw a 2nd ticket at random. Probability is $1/n$ as before
- Q. What is the probability of two ones in a row?
- A. $\text{prob} = \frac{1}{n} \times \frac{1}{n} = \frac{1}{n^2}$ i.e. $1/n$ th of the first $1/n$
- Alternatively, for drawing 2 tickets consecutively as above there are
 - n possibilities for the 1st and
 - n possibilities for the 2nd

Thus there are n^2 pairs - think of this as *a new box of pairs* and each *pair* has the same chance $1/n^2$

This is the uniform distribution again, but on the n^2 pairs.

```
BarChart3D[Table[1/25, {i, 5}, {j, 5}], AxesLabel -> {"x1", "x2", "f(x1, x2)"}];
```

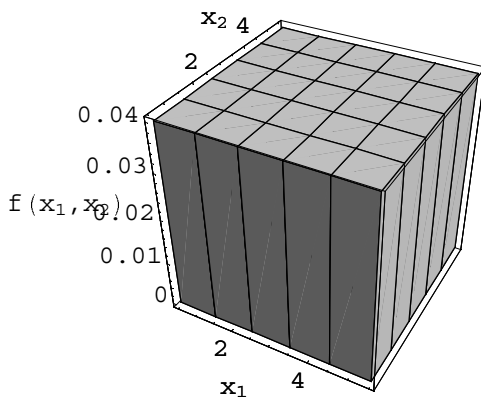


Figure 2

Plot to show the uniform distribution for picking a pair of tickets for $n = 5$. We can represent the random value of the first ticket as X_1 and the random value of the second ticket as X_2 . The possible values taken by these random variables are x_1 and x_2 respectively. The vertical axis is $f(x_1, x_2) = \mathbb{P}(X_1 = x_1, X_2 = x_2)$.

Example 1.3.4 Take the previous example and specialise it to the case $n = 7$.

Q. What is the probability of 2 even numbers?

List all outcomes:

```
Outer[ToString[#{#1, #2}] &, {1, 2, 3, 4, 5, 6, 7}, {1, 2, 3, 4, 5, 6, 7}] // TableForm
Outer[ToString[#{#1, #2}] &, {2, 4, 6}, {2, 4, 6}] // TableForm
```

{1, 1}	{1, 2}	{1, 3}	{1, 4}	{1, 5}	{1, 6}	{1, 7}
{2, 1}	{2, 2}	{2, 3}	{2, 4}	{2, 5}	{2, 6}	{2, 7}
{3, 1}	{3, 2}	{3, 3}	{3, 4}	{3, 5}	{3, 6}	{3, 7}
{4, 1}	{4, 2}	{4, 3}	{4, 4}	{4, 5}	{4, 6}	{4, 7}
{5, 1}	{5, 2}	{5, 3}	{5, 4}	{5, 5}	{5, 6}	{5, 7}
{6, 1}	{6, 2}	{6, 3}	{6, 4}	{6, 5}	{6, 6}	{6, 7}
{7, 1}	{7, 2}	{7, 3}	{7, 4}	{7, 5}	{7, 6}	{7, 7}

Table 1

Table of all possible pairs of tickets drawn consecutively with replacement.

{2, 2}	{2, 4}	{2, 6}
{4, 2}	{4, 4}	{4, 6}
{6, 2}	{6, 4}	{6, 6}

Table 2

Table of all possible pairs of tickets with even numbers drawn consecutively with replacement.

A. Number of pairs is $7^2 = 49$, number of pairs with both numbers even is $3^2 = 9 \Rightarrow$
 [probability of 2 even tickets] = $\frac{9}{49} \approx 18\%$.

```
Show[GraphicsArray[{ListDensityPlot[Table[0, {i, 7}, {j, 7}]], ListDensityPlot[
  Table[If[And[2 IntegerPart[i / 2] == i, 2 IntegerPart[j / 2] == j], 0, 1], {i, 7}, {j, 7}]]]]
```

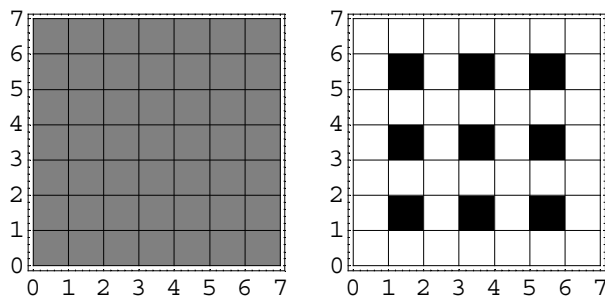


Figure 3

Density plots to show the total number of pairs and the total number of pairs of even tickets, respectively. Tickets are drawn with replacement.

1.3.2 Picking Objects Without Replacement

Example 1.3.5 Again consider a box containing n tickets, numbered $1, 2, \dots, n$.

- Draw a ticket at random (distribution is uniform, but only for those tickets that remain.)
- Do **not** replace into the box
- Draw a 2nd ticket at random. Probability is no longer $1/n$. Most tickets are more probable, one is (much!) less probable.
- Probability of drawing two ones is zero.
- Over all, there are only $n \times (n - 1)$ possibilities.

```
Flatten[Map[Permutations, KSubsets[{1, 2, 3, 4, 5, 6, 7}, 2]], 1]
```

```
{ {1, 2}, {2, 1}, {1, 3}, {3, 1}, {1, 4}, {4, 1}, {1, 5}, {5, 1},
  {1, 6}, {6, 1}, {1, 7}, {7, 1}, {2, 3}, {3, 2}, {2, 4}, {4, 2},
  {2, 5}, {5, 2}, {2, 6}, {6, 2}, {2, 7}, {7, 2}, {3, 4}, {4, 3},
  {3, 5}, {5, 3}, {3, 6}, {6, 3}, {3, 7}, {7, 3}, {4, 5}, {5, 4}, {4, 6},
  {6, 4}, {4, 7}, {7, 4}, {5, 6}, {6, 5}, {5, 7}, {7, 5}, {6, 7}, {7, 6} }
```

```
BarChart3D[Table[If[i ≠ j,  $\frac{1}{42}$ , 0], {i, 7}, {j, 7}], AxesLabel → {"x1", "x2", "f(x1, x2)"}];
```

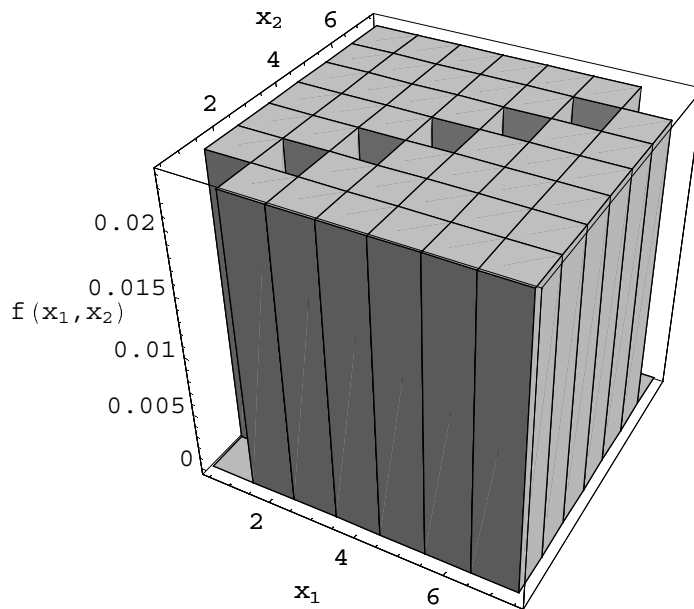


Figure 4

Bar chart of probabilities for pairs of tickets drawn consecutively without replacement.

Q. What is the probability of 2 even numbers?

- Number of possibilities for the first being even is $\#\{2, 4, 6\} = 3$
- Number of possibilities for the 2nd being even is 2 (one even number removed, two remain). (N.B. 2 possibilities for *each* of the 3 from the first draw)

$\Rightarrow 3 \times 2$ possibilities for the event "2 even"

$$\Rightarrow \frac{3 \times 2}{7 \times 6} = \frac{1}{7} \approx 14\%$$

```
Show[GraphicsArray[{ListDensityPlot[Table[If[i ≠ j, 0, 1], {i, 7}, {j, 7}]], ListDensityPlot[
  Table[If[And[i ≠ j, 2 IntegerPart[i/2] = i, 2 IntegerPart[j/2] = j], 0, 1], {i, 7}, {j, 7}]]]}
```

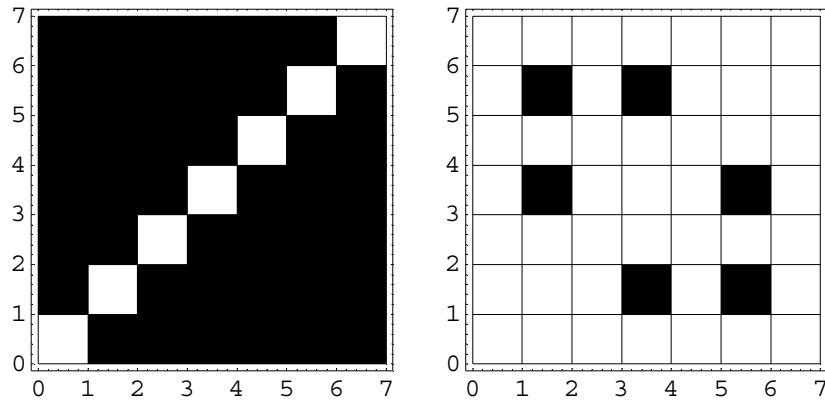


Figure 5

Density plots to show the total number of pairs and the total number of pairs of even tickets, respectively, when tickets are drawn without replacement. The plot on the left has 42 black squares, the plot on the right has 6 black squares.

1.3.3 Brief introduction to *sample spaces*

Definition 1.3.2

The set of all possible outcomes of an experiment is called the **sample space** and is denoted by S

Example 1.3.6 2 (fair) dice are rolled. What is the probability that "the sum equals 3"

There are 36 possible outcomes.

```
Outer[ToString[{-#1, #2}] &, {1, 2, 3, 4, 5, 6}, {1, 2, 3, 4, 5, 6}] // TableForm
```

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

Notation: a more compact way to write this is $S = \{(i, j) \mid i, j = 1, \dots, 6\}$

```
Show[GraphicsArray[{ListDensityPlot[Table[0, {i, 6}, {j, 6}]],
  ListDensityPlot[Table[If[i + j = 3, 0, 1], {i, 6}, {j, 6}]]}]
```

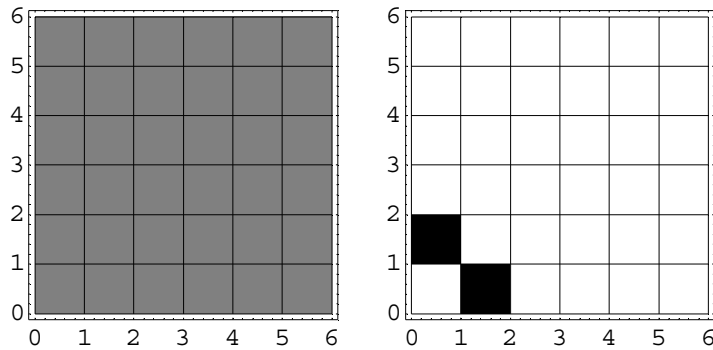


Figure 6

Density plots to show all possible scores on the two dice (left) and those outcomes that sum to 3 (right).

We expect to see (1,2) $\frac{1}{36}$ of the time and (2,1) $\frac{1}{36}$ of the time.

Thus we will have the event "sum equals 3" $\frac{1}{36} + \frac{1}{36} = \frac{1}{18}$ of the time.

$$\mathbb{P}\{\{\text{sum equals 3}\}\} = \frac{1}{18}$$

Notation: a more compact way to write the event that the sum equals 3 is

$$A = \{(i, j) \mid i, j = 1, \dots, n; i + j = 3\}$$

or more concisely

$$A = \{(i, j) \in S; i + j = 3\}$$

Exercise 1.3.1 Find the probability that

- the sum of the two dice is 5
- the difference between the scores on the dice is 1.

(Hint: use the diagrams below.)

```
Show[GraphicsArray[
  Map[ListDensityPlot[Table[If[#, 0, 1], {i, 6}, {j, 6}], DisplayFunction->Identity] &,
  {i != j, Abs[i - j] = 1, i > j, And[i + j == 4, i != j], And[i + j == 5, i != j]}]]];
```

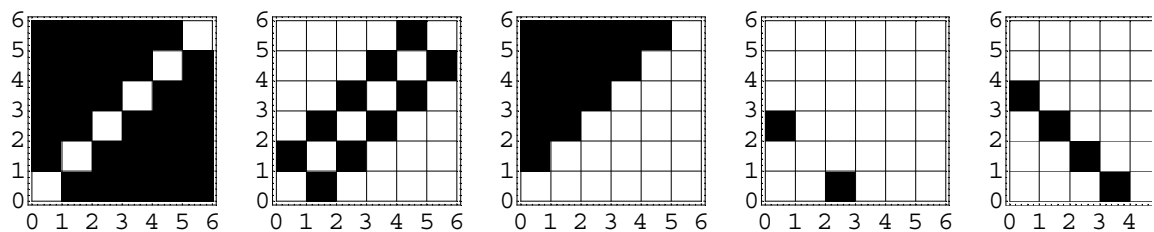


Figure 7

Plots of (a) the sample space, (b) $|i - j| = 1$ (c) $j > i$ (d) $i + j = n$, for the case $n = 4$, (e) $i + j = n$, for the case $n = 5$

1.3.4 A physical definition of probability

- Distinctive feature of games of chance: outcome of a given experiment cannot be predicted with certainty
- Collective results many trials display some regularity
- E.g. probability of "heads" in tossing a coin equals one-half (relative frequency interpretation) \Rightarrow many tosses the relative frequency of "heads" occurs $\sim 1/2$; no implication re outcome of any given toss.

Interpretation 1

probability = the relative frequency that something happens
 = the proportion of the time that it happens

Interpretation 2

(subjective) probability =
 the relative frequency that you *think* something would happen *if* you could repeat many trials

- Thus probability is given as a number between 0 and 1, i.e. $\frac{1}{2}$ of the time, $\frac{1}{4}$ of the time, 0.875 of the time, etc.
- 0 means *never*, 1 means *always*.
- Or as a percentage 50%, 25% etc.

1.3.5 Probability distribution functions

Example 1.3.7 "Suppose ⁷ a ¹ word ⁴ is ² picked ⁶ at ² random ⁶ from ⁴ this ⁴ sentence ⁸"

Find the *distribution* of the length of the word picked.

[picked]				[this]
	[a]			
	[from]			[sentence]
[this]		[random]		
			[at]	[suppose]

Table 3

Experiment

Pick words at random:

```
RandomWord := {"suppose", "a", "word", "is", "picked",
               "at", "random", "from", "this", "sentence"}[[Random[Integer, {1, 10}]]]
```

```
RandomWord
```

```
a
```

Find frequencies for each word:

```
ListOfWords := Table[RandomWord, {n, 1000}];
```

```
ListOfWords
```

```
frequenciesEachWord = Map[Count[ListOfWords, #] &,
                          {"suppose", "a", "word", "is", "picked", "at", "random", "from", "this", "sentence"}];
BarChart[frequenciesEachWord, AxesLabel -> {"word", "freq[word]"}];
```

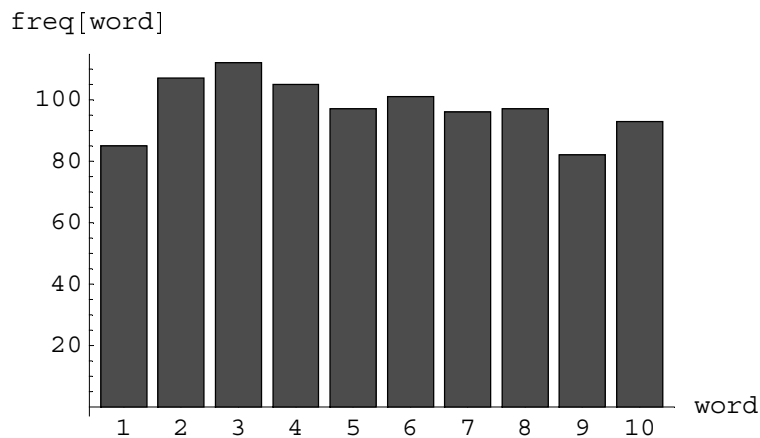


Figure 8

Bar chart to show the frequencies obtained for each of the 10 words in 1000 draws from the sentence. Each "bin" gets ~ 100, i.e. 10% of the 1000 trials

Find frequencies for each word length:

```
Map[StringLength, ListOfWords]
```

```
frequenciesEachWordLength = Table[Count[Map[StringLength, ListOfWords], n], {n, 10}];
BarChart[frequenciesEachWordLength, AxesLabel -> {"word length", "freq[word length]"}];
```

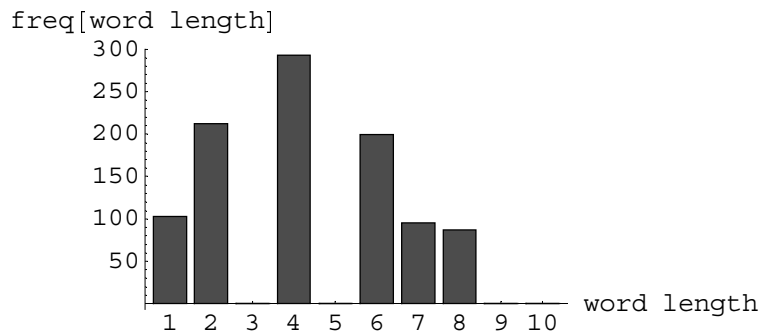


Figure 9

Bar chart to show the frequencies obtained for the word length in 1000 draws. Each "bin" gets a fraction of the 1000 trials in proportion to the probability of that word length arising.

Discussion

Idea: if you reach in and pick out a word at random then 4-letter words are the most probable.

- $\frac{3}{10}$ of the words have 4 letters
- $\frac{2}{10}$ of the words have 6 letters etc.

The number of letters is called a *random variable*.

Let

- X be the (random) number of letters in each word
- x be the possible values taken by X
- $p(x) = \mathbb{P}[X = x]$ is the *probability mass function*

```
TableForm[Transpose[{Table[i, {i, 10}], {1, 2, 0, 3, 0, 2, 1, 1, 0, 0} / 10}],
  TableHeadings -> {None, {"x", "p(x)}}]
```

x	$p(x)$
1	$\frac{1}{10}$
2	$\frac{1}{5}$
3	0
4	$\frac{3}{10}$
5	0
6	$\frac{1}{5}$
7	$\frac{1}{10}$
8	$\frac{1}{10}$
9	0
10	0

Table 4

Probability mass function for X , the number of letters in the word selected at random from the sentence.

```
BarChart[{1, 2, 0, 3, 0, 2, 1, 1, 0, 0} / 10, AxesLabel -> {"x", "P[X = x]"}];
```

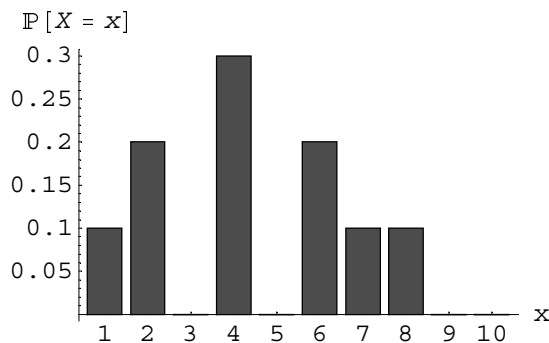


Figure 10

Bar chart of the probability mass function for x , the number of letters in the word selected at random from the sentence.

1.3.6 Great Expectations

Example 1.3.8 "Suppose a word is picked at random from this sentence". Find the average word length.

Take the *dot product* or *inner product* of the possible values for X and their probabilities (you can implement this in Microsoft Excel using the command `SumProduct`).

$$\begin{aligned}
 &1 \times \frac{1}{10} + 2 \times \frac{1}{5} + 3 \times 0 + 4 \times \frac{3}{10} + 5 \times 0 + 6 \times \frac{1}{5} + 7 \times \frac{1}{10} + 8 \times \frac{1}{10} + 9 \times 0 + 10 \times 0 = 4.4 \\
 &= \sum_{x=1}^{10} x p(x)
 \end{aligned}$$

Call this the *expected value* of x

Interpret the expected value as the average value that one would obtain per trial if we performed many trials:

Suppose N trials:

- 4 occurs $\frac{3}{10} N$ times (roughly)
- x occurs $p(x) N$ times (see above - when $N = 1000$, number of words length 4 was ~300)

Total word length in N trials (e.g. = $1 + 4 + 4 + 7 + 2 + 6 + \dots + 1 + 7$)

$$\approx \sum_{x=1}^{10} N p(x) x = \frac{\text{Contribution from 1 letter words}}{N \times p(1) \times 1} + \frac{\text{Contribution from 2 letter words}}{N \times p(2) \times 2} + \dots + N \times p(10) \times 10$$

Average word length (e.g. = $\frac{1+4+4+7+2+6+\dots+1+7}{1000}$)

$$\approx \frac{N \sum_x p(x) x}{N} = \sum_{x=1} p(x) x$$

Experiment

Average is total word length (1000 randomly selected words strung end to end) divided by number of words (1000):

```
AllWordsJoined := Apply[StringJoin, ListOfWords]
```

```
Short[AllWordsJoined, 4]
```

```
fromsentencepickedthiswordawordisrandomath
... ssentencethisafromrandomwordfromisfromword
```

Total word length for this particular set of 1000 words (reevaluate code to repeat the experiment):

```
StringLength[AllWordsJoined]
```

```
4403
```

Average word length for this run of the experiment:

```
StringLength[AllWordsJoined] / 1000 // N
```

```
4.413
```

1.3.7 Birthday Problems

■ 1.3.7.1 Introduction

... just one of the latest manifestations of our natural tendency to read significance into coincidence..

Paulos in *Once Upon a Number*

... misunderstandings of probability may be the greatest of all impediments to scientific literacy.

Stephen Jay Gould

... probability is a subject full of seductively plausible arguments which can fall apart when you examine them

John Haigh in *Taking Chances* P86

- Often used to illustrate the frequently observed fact that in the field of probability, human intuition very often leads the unsuspecting individual to completely the wrong answer. (See the quote from de Moivre at the start of the notes.)
- Coincidences are far more likely to happen than we would imagine.
- See Paulos 1998 P2, P57, for discussions on bible codes: *equidistant letter sequences* "ELS" found in the Torah, in Hebrew including names of holy men; bestseller *The Bible Code*: prophesy of Itzhak Rabin's assassination!

■ 1.3.7.2 The birthday problem

- Probability that in a randomly selected group of n people at least two have the same birthday.
- Probability that any two have the same birthday is $\frac{1}{365}$
- For there to be approximately 50% chance of 2 birthdays in common, sensible guess would be group size $\approx 186 \approx \frac{365}{2} ???$

■ 1.3.7.3 Solution to the birthday problem

- Assume for simplicity that a year contains 365 days and that each day is equally likely to be the birthday of a randomly selected person,
- In a group of n people there are 365^n possible combinations of birthdays (each of the n individuals can have their birthday on any of the 365 days.)
- *Notation:*
 - Let x_i be the birthday of the i th person.

- Outcome/ sample element looks like (x_1, x_2, \dots, x_n) (Each of these n -tuples is equally likely)
- $S = \{(x_1, x_2, \dots, x_n) \mid x_i = 1, \dots, 365; i = 1, \dots, n\}$
- $\#S = 365^n$ (cardinality)

E.g. in the classroom $n = 200$, and number of possible outcomes is $365 \times 365 \times 365 \times 365 \times \dots \times 365 = 365^{200}$

```
365200
```

```
28745700163878677369255751671938997328043321071543993129964932626032691741
5006720920727987856834326888856973735702530622005313743060465103930403110
4043990114027315892197270653798385001138186456600074555842057517424089960
9583715312889675546820839698499284166151954455518688601674962205296149492
4924588979136534367019274710735556307677482007778428858895059612165185584
4199747003074754327472345859123480128024003799711458567097021717958786604
3651685536264472761579962066926050706650030974742549005895853042602539062
5
```

```
ScientificForm[N[365200], 3]
```

```
2.87 × 10512
```

- *Trick*: because probabilities of all possible outcomes sum to 1, we can deduce the probability that there are 2 or 3 or 4 or 5 or any number of birthdays in common up to a total of n , by finding the probability that there are no birthdays in common and subtracting it from one.
- Of these possibilities, how many have all of the n individuals having different birthdays
 - the first person could have his or her birthday on any of the 365 days of the year,
 - the second person could have his or her birthday on any of the other 364 days of the year,
 - the third could have his or her birthday on any of the other 363 days of the year, ...,
 - the n th could have his or her birthday on any of the remaining $365 - n + 1$ days.
- Hence, the number of ways that all n people can have different birthdays is $365 \times 364 \times \dots \times (365 - n + 1)$ (# outcomes with no birthdays in common)

E.g. if $n = 200$, number of outcomes with no bds in common is $365 \times 364 \times \dots \times 166$

```
365! / 165!
```

```
46284185379961942698041843364378018542868496047934948559560091963227316721
7986037953437132273710424628367093476202863726812015445761568036325893200
7898857796106055773174812813124321126959891011796249628590413046291424041
8192901297362047071470425558579611633804185967899311980585801785900694169
9317749897045664722537203595833608212704321274564551260706940515178932308
1366212063089884937871767101789344229300217695621285035214934481305600000
000000000000000000000000000000000000000000000000000000000000000000000000
```

```
ScientificForm[N[365! / 165!], 3]
```

```
4.63 × 10482
```

- Probability that at least two have the same birthday is

$$1 - \frac{365 \times 364 \times \cdots \times (365 - n + 1)}{365^n} = 1 - \frac{365!}{(365 - n)!} \frac{1}{365^n}$$

$$= 1 - {}_{365}P_n \left(\frac{1}{365} \right)^n = 1 - \frac{365^{(n)}}{365^n}.$$

- Produce a table of probabilities for different sized groups:

```
TableForm[Map[{#, NumberForm[N[1 -  $\frac{365!}{(365 - \#)!} \frac{1}{365^\#}$ ], {7, 2}]} &, {5, 18, 23, 32, 42, 58, 100}],
TableHeadings → {None, {"n", "P[#bd≥2]}]}
```

n	P[#bd≥2]
5	0.03
18	0.35
23	0.51
32	0.75
42	0.91
58	0.99
100	1.

Table 5

Table to show how the probability that there are 2 or more birthdays in common within a group of individuals depends on the size of the group.

```
ListPlot[Table[{n, N[1 -  $\frac{365!}{(365 - n)!} \frac{1}{365^n}$ ]}, {n, 1, 100}], PlotJoined → True,
PlotStyle → {Thickness[0.02], Hue[0]}, AxesLabel → {"n", "P[#bd≥2]}];
```

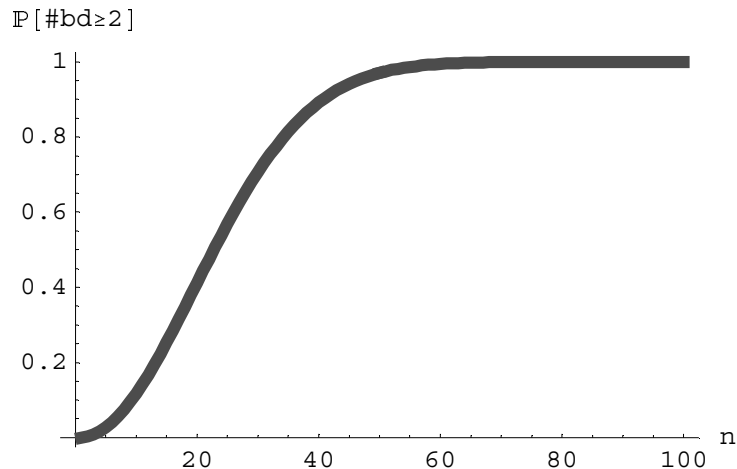


Figure 11

Plot to show how the probability that there are 2 or more birthdays in common within a group of individuals depends on the size of the group.

■ 1.3.7.4 Discussion

- Illustrates that applications of probability theory to the physical world require assumptions and simplifications
 - a year has 365 days (one year in four it is 366)
 - all days are equally likely to be the birthday of a random individual (birthdays are distributed in a non-uniform manner)

The mathematicians motto is that if a problem looks hard, replace it with an easier one.

John Haigh in *Taking Chances* P123

- John Haigh has a plausibility argument for why the size of the group required for 50% chance of 2 or more birthdays in common is as small as it is:
 - Probability that two have the same birthday is $\frac{1}{365}$
 - In a group of 23 people there are ${}_{23}C_2 = 253$ pairs of people. (In a group of 200 people there are ${}_{200}C_2 = 19900$ pairs of people).
 - Naive ratio, $\frac{{}_{23}C_2}{365} = \frac{253}{365} = 0.69$ is not correct, but is of the right order of magnitude (in the right ball park)

■ 1.3.7.5 Approximate solution

For the purposes of finding an approximation formula, we can write the probability of there being no birthdays in common in a group of n people in an alternative form.

$$\begin{aligned} \frac{365 \times 364 \times \cdots \times (365 - n + 1)}{365^n} &= \frac{365}{365} \times \frac{364}{365} \times \frac{363}{365} \times \cdots \times \frac{(365 - n + 1)}{365} \\ &= 1 \times \frac{365 - 1}{365} \times \frac{365 - 2}{365} \times \cdots \times \frac{(365 - (n - 1))}{365} \\ &= \prod_{i=1}^{n-1} \frac{365 - i}{365} \end{aligned}$$

It can be shown (Problem Sheet 1, Question 7) that

$$\prod_{i=1}^{n-1} \frac{365 - i}{365} \approx \exp\left(-\frac{n(n-1)}{730}\right)$$

using the rule for the summation of arithmetic progressions: $\sum_{i=1}^n i = \frac{n(n+1)}{2}$

```
MultipleListPlot[Table[{n, N[1 -  $\frac{365!}{(365 - n)! 365^n}$ ]}, {n, 1, 100}],
Table[{n, 1 - Exp[-n(n - 1) / 730]}, {n, 1, 100}], PlotJoined -> {True, True},
PlotStyle -> {{Hue[0]}, {Hue[0.5]}}, SymbolShape -> {None, None}, AxesLabel -> {"n", "P[#bd ≥ 2]"}];
```

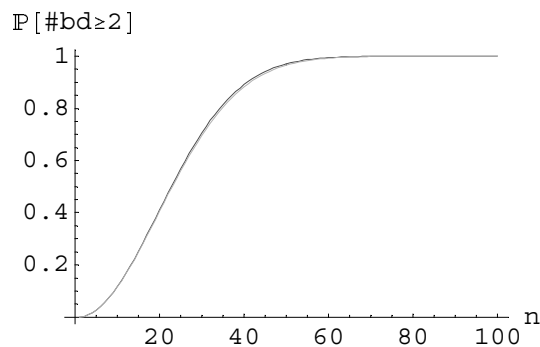


Figure 12

Comparison between exact (red) and approximate (blue) probabilities for 2 or more birthdays in common in a group of size n .

$$\prod_{i=1}^{n-1} \left(1 - \frac{i}{365}\right) \approx e^{-n(n-1)/730}$$

Consider the Taylor expansion of the exponential function $e^{-x} = 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \cdots + (-1)^k \frac{x^k}{k!} + \cdots$, valid for any x .

The truncated Taylor expansion, $e^{-x} \approx 1 - x$, is valid for "sufficiently" small x .

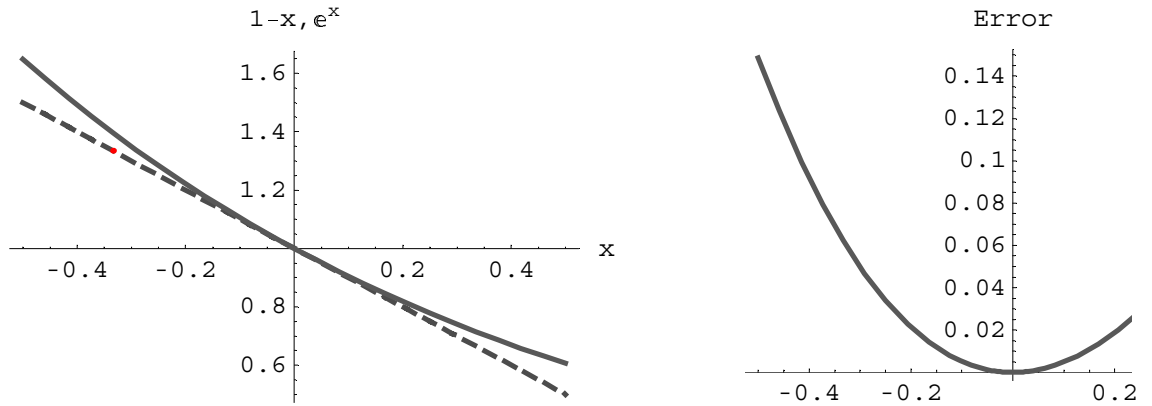
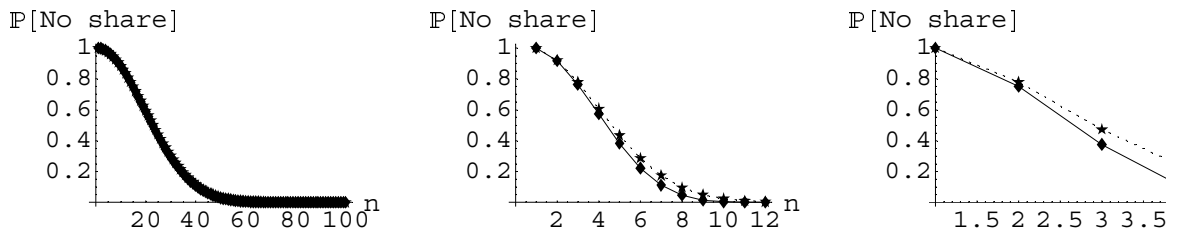


Figure 13

(Left) Comparison of e^{-x} (blue) and $1 - x$ (red, dashed) and (right) difference between e^{-x} and $1 - x$ (the error in the approximation).

Practice problem asked you to consider for which values of n the approximation was adequate:



Comparison of exact (solid) and approximate (dashed) expressions for the probability of there being no birthdays in common for (Left) $n = 365$, (Middle) $n = 12$ and (Right) $n = 4$.

■ 1.3.7.6 Alternative birthday problem

Same context, but *very* different question:

- In a group of 200 randomly selected people, pick one individual, Alice. What is the probability that someone else has her birthday, b , say?
- The set of all possible outcomes to the experiment, called the *sample space* is

$$S = \{(x_1, x_2, \dots, x_{199}) \mid x_i = 1, \dots, 365; i = 1, \dots, 199\}$$

e.g. of an outcome

$$s_1 = \overbrace{\{1, 3, 3, 4, 7, 8, 8, \dots, 320, 322, 364\}}^{199 \text{ birthdays}}$$

- How many outcomes are in S ?

$$\# S = 365^{199}$$

- If we let A denote the event that none of the coordinates x_i correspond to b , then formally

$$A = \{(x_1, x_2, \dots, x_{199}) \in S \mid x_i \neq b \ \forall i\}$$

$$\# A = 364^{199}$$

- *Conclusion:* the probability that at least someone in the group has Alice's birthday is

$$\mathbb{P}[A^c] = 1 - \frac{364^{199}}{365^{199}} = 42 \%$$

- We can generalise this result to a group of size n

Result 1.3.2

$$\mathbb{P}[\text{In a group of size } n, \geq 1 \text{ has birthday } b] = 1 - \frac{364^n}{365^n}$$

Approximation

$$1 - \frac{364^n}{365^n} = 1 - \left(1 - \frac{1}{365}\right)^n \approx 1 - e^{-n/365}$$

```
TableForm[Map[{\#, n /. Flatten[NSolve[1 - (1 - 1/365)^n == \#, n]}, -365 Log[1 - \#]} &,
{0.5, 0.75, 0.95, 0.99}], TableHeadings -> {None, {"P[A^c]", "n"}}]
```

$p = \mathbb{P}[A^c]$	n	$-365 \text{Log}[1 - p]$
0.5	252.7	253.0
0.75	505.3	506.0
0.95	1091.9	1093.4
0.99	1678.6	1680.9

Table 6

Table to show how large n needs to be to reach probabilities for $A^c = \{\text{At least someone has birthday } b\}$ of $\frac{1}{2}$, $\frac{3}{4}$, $\frac{19}{20}$, and $\frac{99}{100}$. The third column contains the approximate value of n from the approximation formula.

On the following plot we compare the original and alternative birthday problems.

```
Plot[{1 - Exp[-n(n-1)/730], 1 - e^{-n/365}}, {n, 1, 500},
PlotStyle -> {{Thickness[0.02], Hue[0]}, {Thickness[0.02], Hue[0.6], Dashing[.03]}},
PlotRange -> All, AxesLabel -> {"n", "P[A^c]"}];
```

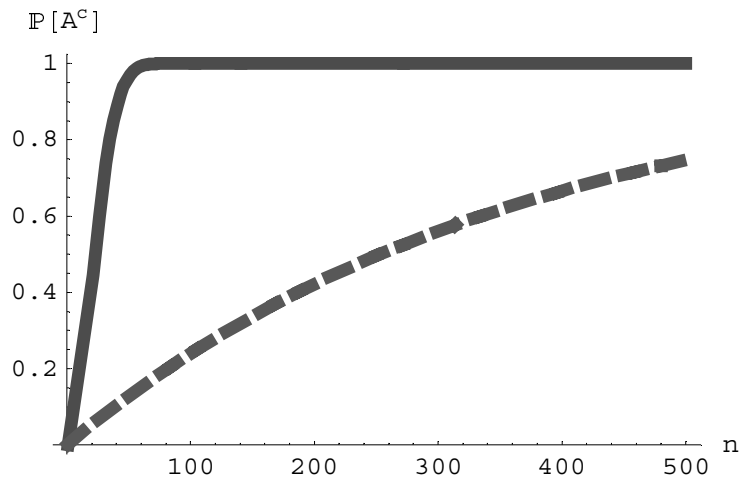


Figure 14

Probability (approximate) that in a group size n

- two or more people share a birthday (red, continuous)
- one or more people have a particular birthday (blue, dashed)

1.3.8 Mathematical Model

There are 3 ingredients to a mathematical model for an experiment involving uncertainty:

- Sample space S
- Events A, B etc. $\subset S$
- Probability $\mathbb{P}[A]$, $\mathbb{P}[B]$, $\mathbb{P}[A \cup B]$... etc.

Example 1.3.9 Toss 2 dice.

- Sample space:
 - $S = \{(i, j) \mid i, j = 1, \dots, 6\}$
 - $\#S = 36$
 - die said to be *fair* if there is a uniform distribution on these points
- Events are various subsets of S . Consider
 - $A = \{(i, j) \in S \mid i + j = 4\} = \{(1, 3), (2, 2), (3, 1)\}$
 - $B = \{(i, j) \in S \mid \max(i, j) = 3\} = \{(1, 3), (2, 3), (3, 3), (3, 2), (3, 1)\}$
- Number of outcomes in each event:
 - $\#A = 3$
 - $\#B = 5$

- Probabilities of each event

- $\mathbb{P}[A] = \frac{3}{36}$

- $\mathbb{P}[B] = \frac{5}{36}$

```
Show[GraphicsArray[
  {ListDensityPlot[Table[If[i + j == 4, 0, 1], {i, 6}, {j, 6}], AxesLabel -> {"i", "j"}],
  ListDensityPlot[Table[If[Max[i, j] == 3, 0, 1], {i, 6}, {j, 6}], AxesLabel -> {"i", "j"}]}];
```

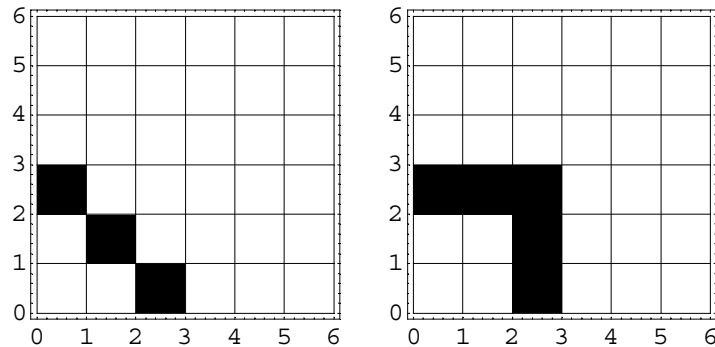


Figure 15

Events $A = \{(i, j) \in S \mid \max(i, j) = 3\}$ and $B = \{(i, j) \in S \mid \max(i, j) = i + j = 4\}$, displayed in the form of a density plot. Outcomes in black are in the event.

- Events combined using set operations:

- $A \cup B = \{\text{total is 4 or maximum is 3}\}$

- $A \cap B = \{\text{total is 4 and maximum is 3}\}$

- Number of outcomes

- $\#A \cup B = 6$

- $\#A \cap B = 2$

- Probabilities of each event

- $\mathbb{P}[A \cup B] = \frac{6}{36}$ ← can use $\mathbb{P}[A] = \frac{\#A}{\#S}$ for finite, equiprobable s. spaces.

- $\mathbb{P}[A \cap B] = \frac{2}{36}$

Probability in finite, equiprobable sample spaces

$$\mathbb{P}[A] = \frac{\text{number of outcomes in } A}{\text{number of outcomes in } S} = \frac{\#A}{\#S}$$

```
Show[GraphicsArray[{ListDensityPlot[Table[If[Or[i + j == 4, Max[i, j] == 3], 0, 1], {i, 6}, {j, 6}],
  AxesLabel -> {"i", "j"}], ListDensityPlot[
  Table[If[And[i + j == 4, Max[i, j] == 3], 0, 1], {i, 6}, {j, 6}], AxesLabel -> {"i", "j"}]}];
```

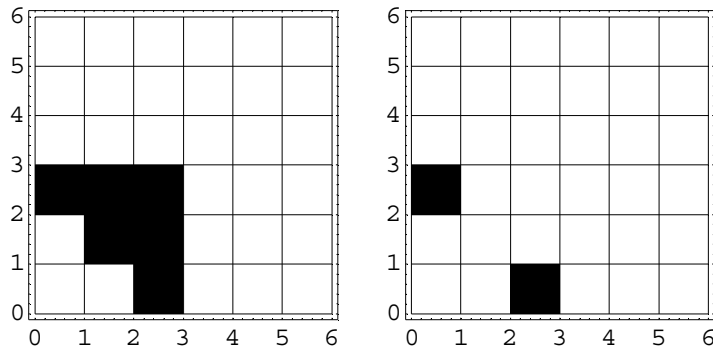


Figure 16

Events $A \cup B$ and $A \cap B$, displayed in the form of a density plot. Outcomes in black are in the event.

N.B. Not necessarily true that $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B]$ (in our example, $\frac{6}{36} \neq \frac{3}{36} + \frac{5}{36}$)

Theorem 1.3.3

$\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B]$ if A and B are *disjoint* (*mutually exclusive*)

- We could introduce a 3rd event, $C = \{2 \text{ nd die shows a } 3\} = \{(i, j) \in S \mid j = 3\}$. Then
 - $A \cup B \cup C = \{\text{all outcomes in any of } A, B \text{ or } C\}$
 - $A \cap B \cap C = \{\text{all outcomes in all of } A, B \text{ and } C\}$

Example 1.3.10 Pick a card from a pack of 52 playing cards. Let A be the event of picking a heart, B be the event of picking a picture card (any of the royalty).

Some typical outcomes are $5\clubsuit, 8\diamond$ etc.

$$\#S = 13 \times 4 = 52$$

$$A \cap B = \{J\heartsuit, Q\heartsuit, K\heartsuit\}$$

Some outcomes in $A \cup B$ are $\{J\heartsuit, Q\heartsuit, K\heartsuit, J\clubsuit, Q\clubsuit, K\clubsuit, J\diamond, Q\diamond, K\diamond, A\heartsuit, 2\heartsuit, 7\heartsuit, \dots\}$

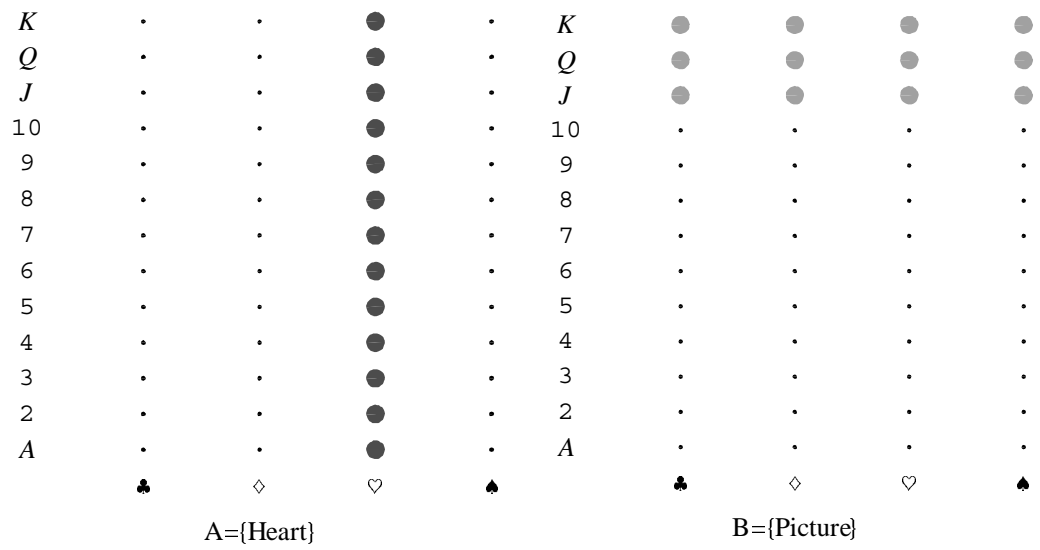


Figure 17

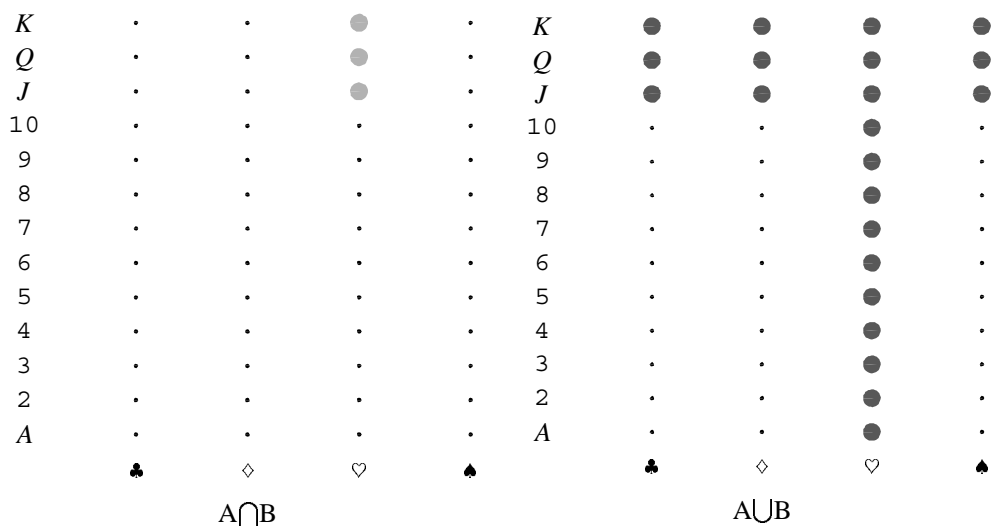


Figure 18

The End

1.3.9 Summary

1.3.1 Picking Objects With Replacement

- Box containing n numbered tickets. Do replace.
 - Probability of "1" is $1/n$ - *uniform* distribution
 - Q. Probability of two ones in a row?
 - A. $\text{prob} = \frac{1}{n} \times \frac{1}{n} = \frac{1}{n^2}$

Thus there are n^2 pairs - and each *pair* has the same chance $1/n^2$

- Case $n = 7$.

Q. What is the probability of 2 even numbers?

A. Number of pairs is $7^2 = 49$, number of pairs with both numbers even is $3^2 = 9 \Rightarrow$
 [probability of 2 even tickets] = $\frac{9}{49} \approx 18\%$.

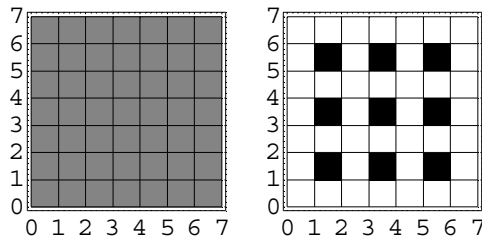


Figure 19

Density plots to show the total number of pairs and the total number of pairs of even tickets, respectively. Tickets are drawn with replacement.

1.3.2 Picking Objects Without Replacement

Box containing n tickets

- Do **not** replace into the box
- Prob. of two of any numbers the same is zero.
- Over all, there are only $n \times (n - 1)$ possibilities.
- Probability of a particular pair $(i, j) = \begin{cases} 0 & i = j \\ \frac{1}{n(n-1)} & i \neq j \end{cases}$

1.3.3 Brief introduction to *sample spaces*

Definition: The set of all possible outcomes of an experiment is called the *sample space*, S

Example 2 (fair) dice. Prob that "the sum equals 3"?

There are 36 possible outcomes. We can list these in tabular form or specify the s.s. by its properties

$$S = \{(i, j) \mid i, j = 1, \dots, 6\} \text{ and the event } A = \{(i, j) \in S; i + j = 3\}$$

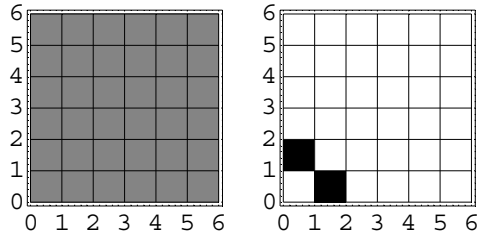


Figure 20

Density plots to show all possible scores on the two dice (left) and those outcomes that sum to 3 (right).

Add the probabilities of each outcome: $\frac{1}{36} + \frac{1}{36} = \frac{1}{18}$

$$P[\{\text{sum equals 3}\}] = \frac{\# A}{\# S} = \frac{1}{18}$$

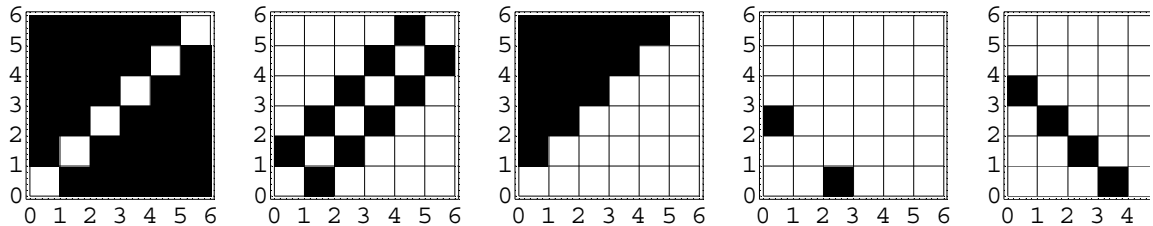


Figure 21

Plots of (a) the sample space, (b) $|i - j| = 1$ (c) $j > i$ (d) $i + j = n$, for the case $n = 4$, (e) $i + j = n$, for the case $n = 5$

1.3.4 A physical definition of probability

probability = the *relative frequency* that something happens

(subjective) probability =

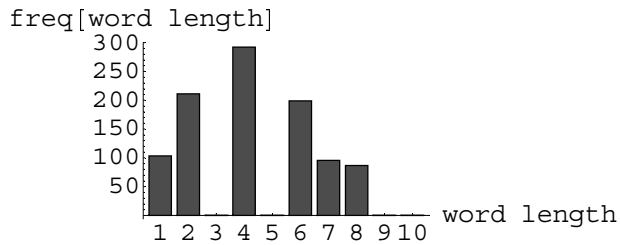
the relative frequency that you *think* something would happen *if* you could repeat many trials

1.3.5 Probability distribution functions

Example "Suppose a word is picked at random from this sentence"
⁷ ¹ ⁴ ² ⁶ ² ⁶ ⁴ ⁴ ⁸

Find the *distribution* of the length of the word picked.

Frequencies from 1000 trials:



4-letter words are the most probable.

The number of letters → a *random variable*.

• $p(x) = \mathbb{P}[X = x]$ is the *probability mass function*

value of X	1	2	3	4	5	6	7	8
10prob	1	2	0	3	0	2	1	1

Probability mass function for X, word length of random word

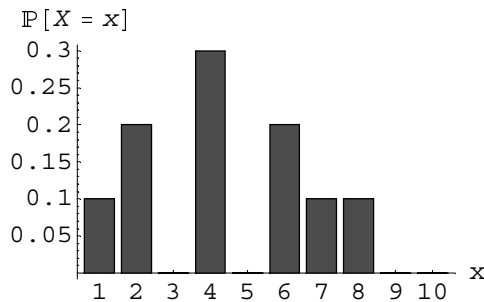


Figure 22

Bar chart of the probability mass function for x, the number of letters in the word selected at random from the sentence.

1.3.6 Expectations

Example "Suppose a word is picked at random from this sentence". Find the average word length.

$$\begin{aligned}
 & 1 \times \frac{1}{10} + 2 \times \frac{1}{5} + 3 \times 0 + 4 \times \frac{3}{10} + 5 \times 0 + 6 \times \frac{1}{5} + 7 \times \frac{1}{10} + 8 \times \frac{1}{10} + 9 \times 0 + 10 \times 0 = 4.4 \\
 & = \sum_{x=1}^{10} x p(x)
 \end{aligned}$$

... the *expected value* of X

Expected value ~ the average value per trial if we performed many trials

$$\text{Total word length in } N \text{ trials} \approx \sum N p(x) x = \frac{\sim \text{Freq of 1 letter words}}{N \times p(1)} \times 1 + \frac{\sim \text{2 letter words}}{N \times p(2) \times 2} + \dots + N \times p(10) \times 10$$

$$\text{Average word length} \approx \frac{N \sum p(x) x}{N} = \sum p(x) x \quad \text{Expected word length}$$

1.3.7 Birthday Problems

The birthday problem - \mathbb{P} that in a group of n people at least two have the same birthday.

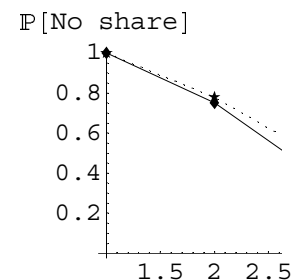
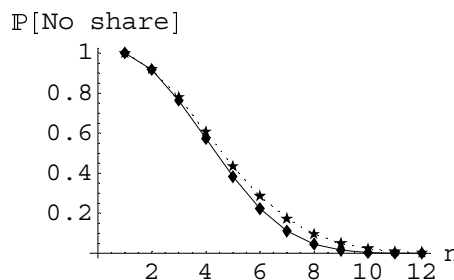
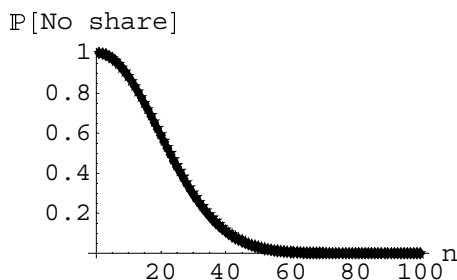
- Assume: year has 365 days and each =ly likely to be the bd of a randomly sel'd person,
- 365^n possible configurations of bds
- $S = \{(x_1, x_2, \dots, x_n) \mid x_i = 1, \dots, 365; i = 1, \dots, n\}$
 $\#S = 365^n$ (cardinality)
- *Trick:* $\mathbb{P}[2 \text{ or more bds}] = 1 - \mathbb{P}[\text{no bds in common}]$
- # ways that all n people can have different birthdays is $365 \times 364 \times \dots \times (365 - n + 1)$
- \mathbb{P} that at least two have the same bd is

$$\mathbb{P}[2 \text{ or more have same bd}] = 1 - \frac{365 \times 364 \times \dots \times (365 - n + 1)}{365^n} = 1 - \frac{365^{(n)}}{365^n}.$$

Approximate solution

$$\frac{365^{(n)}}{365^n} = \prod_{i=1}^{n-1} \frac{365 - i}{365} \approx \exp\left(-\frac{n(n-1)}{730}\right)$$

Use rule for the summⁿ of arithmetic progressⁿ: $\sum_{i=1}^n i = \frac{n(n+1)}{2}$



Comparison of exact (solid) and approximate (dashed) expressions for the probability of there being no birthdays in common for (Left) $n = 365$, (Middle) $n = 12$ and (Right) $n = 4$.

Alternative bd problem

$$S = \{(x_1, x_2, \dots, x_n) \mid x_i = 1, \dots, 365; i = 1, \dots, n\}$$

$$\# S = 365^n$$

- A is the event that none of the coords x_i same as b , then

$$A = \{(x_1, x_2, \dots, x_n) \in S \mid x_i \neq b \ \forall \ i\}$$

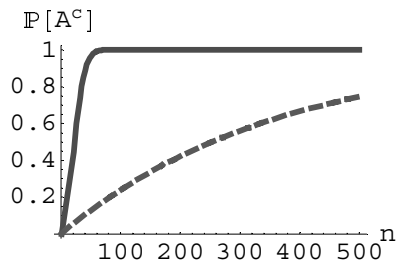
$$\# A = 364^n$$

- *Conclusion*: the probability that at least someone in the group has Alice's birthday is

$$\begin{aligned} \mathbb{P}[\geq 1 \text{ people share bd } b] &= 1 - \frac{364^n}{365^n} \\ &= 42\% \text{ for } n = 199 \end{aligned}$$

Approximation

$$1 - \frac{364^n}{365^n} = 1 - \left(1 - \frac{1}{365}\right)^n \approx 1 - e^{-n/365}$$



1.3.8 Mathematical Model

3 ingredients to math'l model for an exp't involving uncertainty:

- Sample space S
- Events A, B etc. $\subset S$
- Probability $\mathbb{P}[A], \mathbb{P}[B], \mathbb{P}[A \cup B]$... etc.

Example Toss 2 dice.

- $S = \{(i, j) \mid i, j = 1, \dots, 6\}$ $\# S = 36$

- Events are subsets of S .

$$A = \{(i, j) \in S \mid i + j = 4\}$$

$$\# A = 3$$

$$B = \{(i, j) \in S \mid \max(i, j) = 3\}$$

$$\# B = 5$$

● Probabilities of each event

• $\mathbb{P}[A] = \frac{3}{36}$

• $\mathbb{P}[B] = \frac{5}{36}$

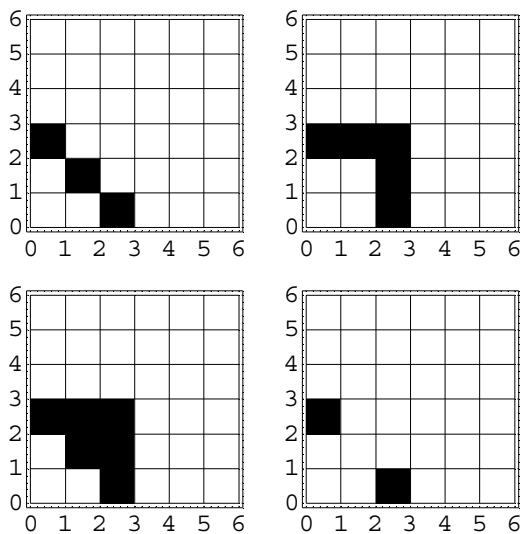
• $A \cup B = \{\text{total is 4 or maximum is 3}\}$ • $\#A \cup B = 6$

• $A \cap B = 6A \cap B = \{\text{total is 4 and maximum is 3}\}$ • $\#A \cap B = 2$

● \mathbb{P}_S

• $\mathbb{P}[A \cup B] = \frac{6}{36}$

• $\mathbb{P}[A \cap B] = \frac{2}{36}$



Events $A = \{(i, j) \in S \mid \max(i, j) = 3\}$ and $B = \{(i, j) \in S \mid \max(i, j) = i + j = 4\}$, displayed in the form of a density plot. Outcomes in black are in the event.

Events $A \cup B$ and $A \cap B$, displayed in the form of a density plot. Outcomes in black are in the event.

Probability in finite, equiprobable sample spaces

$$\mathbb{P}[A] = \frac{\text{number of outcomes in } A}{\text{number of outcomes in } S} = \frac{\#A}{\#S}$$

Theorem $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B]$ if A and B are *disjoint* (events *mutually exclusive*)

• $A \cup B \cup C = \{\text{outcomes in either } A \text{ or } B \text{ or } C\}$

• $A \cap B \cap C = \{\text{outcomes in all of } A, B \text{ and } C\}$

The End

- Miscellaneous Mathematica commands

1.4. Bibliography