Probability for Computer Scientists

This material is provided for the educational use of students in CSE2400 at FIT. No further use or reproduction is permitted.

G.A.Marin, 2008,
All rights reserved.
What is Probability?

- Probability theory (applied) is the study of random outcomes (events) and methods for determining how likely they are to occur.

- A probability is a number that represents how likely is the occurrence of a particular random outcome (event).
  - A probability is a number between 0 and 1.
  - 0 implies that there is no measurable chance the outcome will occur (but it still might!).
  - 1 implies that there is no measurable chance that another outcome will occur.
We use probability all the time.

- The probability that a hurricane will hit Melbourne on Saturday...
- The probability that Dr. Marin will win the lottery this year...
- The probability that 3 students will eat lunch at the Clemente Center (careful!)
- The probability that the S&P 500 index will increase at least 30% this year...
- The probability that you will exhaust your retirement funds before you die...
- The probability that your family will make money on your life insurance policy
Probability Continued

- The probability that the USN will detect a Russian submarine less than 200 miles from the US east coast this week...
- The probability that a missile will strike within 10 meters of its target.
- The probability that a new medication will lower your blood pressure...
- The probability that a roll of a single die will result in a 3...
- The probability that a toss of two coins will result in one head and one tail.
And Statistics?

- In studying probability we look at systems, processes, experiments that produce random (not fixed) outcomes; by studying how the system (etc.) works we attempt to obtain equations/models/tables that show how likely are particular outcomes.

- In statistics we have, say, a database that contains perhaps thousand or millions of outcomes; by studying patterns, averages, or summaries of the data only (mostly) we attempt to obtain equations/models/tables that show how likely are particular outcomes.
For Today’s Graduate, Just One Word: Statistics

By STEVE LOHR
Published: August 5, 2009

MOUNTAIN VIEW, Calif. — At Harvard, Carrie Grimes majored in anthropology and archaeology and ventured to places like Honduras, where she studied Mayan settlement patterns by mapping where artifacts were found. But she was drawn to what she calls “all the computer and math stuff” that was part of the job.

“People think of field archaeology as Indiana Jones, but much of what you really do is data analysis,” she said.

Now Ms. Grimes does a different kind of digging. She works at Google, where she uses statistical analysis of mounds of data to come up with ways to improve its search engine.

Ms. Grimes is an Internet-age statistician, one of many who are changing the image of the profession as a place for dronish number nerds. They are finding themselves increasingly in demand — and even cool.

“I keep saying that the sexy job in the next 10 years will be statisticians,” said Hal Varian, chief economist at Google. “And I’m not kidding.”
The 10 Best Jobs You Can Get Today

Jump to Results

--by Tony Lee

Remember that kid in elementary school who always had a pencil and calculator nearby, and while the rest of us drew pictures, read comic books or played cards, that kid was happily crunching numbers -- for fun. Fast forward 20 years or so, and it turns out that kid probably has one of best careers around today, according to an exclusive new study of the nation's best and worst jobs.

Compiling research on 200 different positions, this year's JobsRated.com report ranks mathematician as the country's best job, followed by actuary and statistician -- three jobs for which a calculator and solitude are prerequisites. On the opposite end of the spectrum, the Monty Python troupe made famous the song, "I'm a lumberjack and I'm OK." Unfortunately, our study finds that lumberjacks have the nation's worst job, followed by dairy farmers and taxi drivers, which seems to bear out the old grade-school adage that "it's better to earn a living with your head rather than with your hands."

Of course, it doesn't take much effort to determine that mathematician is a more attractive job to most people than lumberjack. But ranking 200 jobs from best to worst is no easy feat. To compile this year's report, researchers relied on five criteria to compare jobs as different as librarian and sheet metal worker. Those criteria: stress, physical demands, hiring outlook, compensation and work environment (for more info on scoring, visit our Methodology Page).

If advanced equations aren't your strong suit, however, there are plenty of other jobs that score well, too. After the top three math-oriented careers, the rest of the top 10 read like a who's who of well-educated professions:

Find This Job
View Ratings

1. Mathematician

Applies mathematical theories and formulas to teach or solve problems in a business, educational, or industrial climate.

Find This Job
View Ratings

2. Actuary

Interprets statistics to determine probabilities of accidents, sickness, and death, and loss of property from theft and natural disasters.

Find This Job
View Ratings

3. Statistician

Tabulates, analyzes, and interprets the numeric results of experiments and surveys.

Find This Job
View Ratings
Motivation: Here are some things I’ve done with prob/stat in my career:

- Helped Texaco analyze seismic data to find new oil reserves.
- Helped Texaco analyze the bidding process for offshore oil leases.
- Helped US Navy search for submarines and develop ASW tactics.
- Helped US Navy evaluate capabilities of future forces.
- Helped IBM analyze performance of computer networks in health, airline reservation, banking, and other industries.
- Evaluated new networking technologies such as IBM’s Token Ring, ATM, Frame Relay...
- Analyzed meteor burst communications for US government (and weather and earthquakes...)

Applied Statistics: Probability
Motivation:

- I never could have imagined the opportunities I would have to use this material when I took a course like this one. You can’t either.
- Mostly I cannot teach about all of these problem areas because you must “master” the basics (this course) first. It is the key.
- This takes steady work. Have in mind about 6-8 hours per week (beginning this week!).
  - 20 hours after 3 weeks is nowhere as good.
- If you begin now and learn as you go, you’ll likely succeed. If you don’t do anything, say, until the night before a test or the day before an assignment is due, you will not be likely to succeed.
- Now is the time to commit.
Suggestions on how to study

- Do NOT bring a computer to class. It distracts you.
- Watch, listen and take notes the old fashioned way (pencil and paper). For example, write: slide 7 geom series?? Writing notes/questions actually helps you pay attention.
- As soon after class as possible review the slides covered in class and your class notes. Do not go to the “next” slide until you understand everything on the current slide.
  - Can’t understand something? Ask next class period or come to my office hours. In the words of Jerry McGuire: “Help me to help you.”
  - Answer all points you raised in your notes. Look things up on the web.
- Read the material in the text that corresponds to the slides covered. The syllabus gives you a good indication of where to look.
- Work a few homework problems each night (when homework is assigned).
  - Can’t work one or two of these problems? Come to my office hours. “Help me to help you.”
  - It is much, much worse to come two days before 15 problems are due and say “I can’t do any of these.” I will help you with some of the work, but it’s like trying to learn a musical instrument just by watching me play it.)

Applied Statistics: Probability 1-10
Introduction to Mathcad

- Mathcad is computational software by Mathsoft Engineering and Education, Inc.
- It includes a computational engine, a word processor, and graphing tools.
- You enter equations “almost” as you would write them on paper. They evaluate “instantly.”
- You **MUST** use Mathcad for all homework in this class. Send your mathcad worksheet to gmarin@fit.edu BEFORE class time on the due date. “I could not get a Mathcad terminal” is NO EXCUSE.
  - Name the worksheet HomeworknFirstnameLastname.xmcd
  - The letter n represents the number of the assignment.
  - The “type” .xmcd is assigned automatically.
  - If you do not follow the instructions, you will receive a zero for the homework assignment.
- Mathcad is available (at no charge) in Olin EC 272.
- Tutorial and help are included with the software; resources are also available at mathcad.com.
- Mathcad OVERVIEW is next.
- Note: Homework 1 is on the web site.
Review: Finite Sums of Constants

\[ \sum_{i=1}^{n} 1 = n \quad \sum_{i=0}^{n} 1 = n + 1 \quad \sum_{i=0}^{n-1} 1 = n \]

\[ \sum_{i=1}^{n} c = cn \quad \sum_{i=0}^{n} c = c(n + 1) \quad \sum_{i=0}^{n-1} c = cn, \text{ where } c \text{ is any constant.} \]

Example:
\[ \sum_{i=1}^{100} 9 = 900. \]
Review: Gauss Sum

When Carl Freidrich Gauss, 1777 – 1855, was in the 3rd grade he was asked to compute the sum 1+2+3+...+100. He quickly obtained the following formula:

\[ \sum_{k=1}^{100} k = \frac{100 \cdot 101}{2}. \]

Young Gauss reasoned that if he wrote two of the desired sums as follows:

\[ S = 1 + 2 + 3 + \ldots + 99 + 100 \]
\[ S = 100 + 99 + 98 + \ldots + 1 + 1 \]

then clearly

\[ 2S = \sum_{i=1}^{100} 101 = 100(101) \quad \text{and} \quad S = \frac{100(101)}{2} = 5,050. \]

In general,

\[ \sum_{k=1}^{n} k = \frac{n(n+1)}{2}. \]
Definition of Infinite Series Sum

Begin with the series $S = \sum_{i=1}^{\infty} a_i$ and define the partial sums $S_n = \sum_{i=1}^{n} a_i$, for $n = 1, 2, \ldots$.

The series $S$ is said to converge to $s$ if $\lim_{n \to \infty} S_n = s$.

Example: The geometric series $\sum_{i=0}^{\infty} p^i$ has the partial sums $S_n = \sum_{i=1}^{n} p^i$, for $n = 1, 2, \ldots$.

By division we know that $\frac{1-p^{n+1}}{1-p} = 1 + p + p^2 + \ldots + p^n = S_n$.

By definition the sum of the infinite geometric series is $\lim_{n \to \infty} S_n = \lim_{n \to \infty} \frac{1-p^{n+1}}{1-p}$.

If $|p| < 1$, then the above limit is $\frac{1}{1-p}$.
Review: Gauss Related Series

We obtain an infinite "series" when we wish to "add" infinitely many values. From the sum of constants we might consider the sum $\sum_{k=0}^{\infty} 1$. From calculus we know that we cannot obtain a finite sum from this series and that it "diverges" to $\infty$. However, if we multiply each term by $z^k$ we have the "geometric" series $\sum_{k=0}^{\infty} z^k = \frac{1}{1-z}$ provided $|z| < 1$.

Take the derivative of both sides to obtain the Gauss series

$$\sum_{k=1}^{\infty} k z^{k-1} = \frac{1}{(1-z)^2}, \text{ for } |z| < 1.$$
Try these

\[(a) \sum_{i=0}^{10} \left( \frac{1}{2} \right)^i \]

\[(b) \sum_{i=2}^{10} \left( \frac{1}{2} \right)^i \]

\[(c) \sum_{i=0}^{\infty} \left( \frac{1}{2} \right)^i \]

\[(d) \sum_{i=5}^{\infty} \left( \frac{1}{2} \right)^i \]
Power Series

**Definition:** A power series is a function of a variable (say \( z \)) and has the form

\[
P(z) = \sum_{i=0}^{\infty} c_i (z - a)^i,
\]
where the \( c_i \) are constants and \( a \) is also a constant. In this form the series is said to be "centered" at the value \( a \).

**Convergence:** One of the following is true:

Either (1) the power series converges only for \( z = a \),

Or (2) the power series converges absolutely for all \( z \),

Or (3) there is a positive number \( R \) such that the series converges absolutely for \( |z - a| < R \) and diverges for \( |z - a| > R \). \( R \) is called the radius of convergence.
Ratio Test (Good tool for finding radius of convergence.)

Consider the series \( \sum_{k=0}^{\infty} a_k \) and suppose that \( \lim_{k \to \infty} \frac{|a_{k+1}|}{|a_k|} = \rho > 0 \) (or infinite). If \( \rho < 1 \), then the series converges absolutely. If \( \rho > 1 \) or \( \rho = \infty \), the series diverges.

Example: Determine where the series \( \sum_{k=1}^{\infty} \left( \frac{1}{k} \right) \left( \frac{x}{3} \right)^k \) converges.

The ratio \( \left| \frac{a_{k+1}}{a_k} \right| = \left| \frac{1}{(k+1)} \left( \frac{x}{3} \right)^k \right| \) is \( \frac{k}{k+1} \left| \frac{x}{3} \right| \). The limit \( \lim_{k \to \infty} \frac{k}{k+1} \left| \frac{x}{3} \right| = \left| \frac{x}{3} \right| \); thus, \( \rho = \left| \frac{x}{3} \right| \). It follows that the series converges for \( \left| \frac{x}{3} \right| < 1 \Rightarrow |x| < 3 \), and the series diverges for \( |x| > 3 \).
Review: Other Useful Series

- Exponential function:
  \[ e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \]

- The sum of squares:
  \[ \sum_{k=1}^{n} k^2 = \frac{n(n + 1)(2n + 1)}{6} \]

- Binomial sum:
  \[ \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = (x + y)^n \]

- Geometric sum:
  \[ \sum_{k=0}^{n} z^k = \frac{1 - z^{n+1}}{1 - z} \]
Review: Binomial Coefficients

Pascal’s triangle:

\[
\begin{array}{c|c}
 n & \binom{n}{k} \\

\hline
 0 & 1 \\
 1 & 11 \\
 2 & 121 \\
 3 & 1331 \\
 4 & 14641 \\
 \cdots & \cdots \\
\end{array}
\]

\[
(x + y)^3 = \binom{3}{0}x^3 + \binom{3}{1}x^2y + \binom{3}{2}xy^2 + \binom{3}{3}y^3 = x^3 + 3x^2y + 3xy^2 + y^3
\]

where

\[
\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n^k}{k!} = \frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots1}
\]
Floors and Ceilings

As usual we define

\[ \lfloor x \rfloor = \text{the greatest integer less than or equal to } x \text{ where } x \text{ is any real number}. \]

Similarly we define

\[ \lceil x \rceil = \text{the smallest integer greater than or equal to } x. \]

Thus, \( \lfloor 2.3 \rfloor = 2 \) while \( \lceil 2.3 \rceil = 3. \)

Also, \( \lfloor -7.6 \rfloor = -8 \) while \( \lceil -7.6 \rceil = -7. \)

Only when \( x \) is an integer do we have equality of these functions:

\[ \lfloor x \rfloor = x = \lceil x \rceil. \]

In fact, when \( x \) is not an integer, then \( \lceil x \rceil - \lfloor x \rfloor = 1. \)
Logarithms

Notations:
\[
\begin{align*}
\log n &= \log_2 n \quad \text{(binary logarithm)}, \\
\ln n &= \log_e n \quad \text{(natural logarithm)}, \\
\log^k n &= (\log n)^k \quad \text{(exponentiation)}, \\
\log\log n &= \log(\log n) \quad \text{(composition)}.
\end{align*}
\]

Logarithm functions apply only to the next term in the formula, so that \(\log n + k\) means \((\log n) + k\), and not \(\log(n + k)\).

In the expression \(\log_b a\):

- If we hold \(b\) constant, then the expression is strictly increasing as \(a\) increases.
- If we hold \(a\) constant, then the expression is strictly decreasing as \(b\) increases.

Conversion formula: \[
\log_b a = \frac{\log_{10} a}{\log_{10} b}.
\]

Any base may be used in place of 10.
Log Function (Base 10)
Log(50) Base x

LOGB (x)
Logarithmic Identities

Useful identities for all real $a > 0$, $b > 0$, $c > 0$, and $n$, and where logarithm bases are not 1:

$$a = b^{\log_b a},$$

$$\log_c(ab) = \log_c a + \log_c b,$$

$$\log_b a^n = n \log_b a,$$

$$\log_b a = \frac{\log_c a}{\log_c b},$$

$$\log_b(1/a) = -\log_b a,$$

$$\log_b a = \frac{1}{\log_a b},$$

$$a^{\log_b c} = c^{\log_b a}.$$

Changing the base of a logarithm from one constant to another only changes the value by a constant factor, so we usually don’t worry about logarithm bases in asymptotic notation. Convention is to use $\lg$ within asymptotic notation, unless the base actually matters.
Learning math in many ways is like learning a new language. You may understand a concept but be unable to write it down. You may not be able to read notation that represents a simple concept.

The only way to tackle this “language barrier” is through repetition until the notation begins to look “friendly.”

You must try writing results yourself (begin with blank paper) until you can recall the notation that expresses the concepts.

- Copy theorems from class or from the text until you can rewrite them without looking at the material and you understand what you are writing.

Probability and statistics introduce their own notation to be mastered.

If you begin tonight, and spend time after every class, you can learn this new language.

If you wait until the week before the test, it will be much like trying to learn Spanish, or French, or English in just a few nights (1 night??). It really can’t be done.
Math sentences

Math problems, theorem, solutions ... must be written in sentences that make sense when you read them (EVEN WHEN EQUATIONS ARE USED). You will notice that I am careful to do this to the greatest extent possible on these slides (even knowing that I will explain them in class).

My observation is that most students have no idea how to do this. I often see solutions like the following on homework or test papers:

Evaluate $\int_{0}^{4} x^2 \, dx$. The student writes something like

\[
\begin{align*}
&= x^2 \\
&= \frac{x^3}{3} \\
&= 64 \\
&= \frac{64}{3}.
\end{align*}
\]

The answer is right BUT every step is nonsense.

The = sign means “is equal to.” In the first step, we don’t know to what the equality refers. The equality is simply wrong in the next two steps.

Instead write:

\[
\int_{0}^{4} x^2 \, dx = \frac{x^3}{3} \bigg|_{0}^{4} = \frac{64}{3} - 0 = \frac{64}{3}.
\]
**Permutations and Combinations**

Suppose that we have $n$ objects $O_1, O_2, \ldots, O_n$.

A permutation of order $k$ is an "ordered" selection of $k$ of these for $1 \leq k \leq n$.

A combination of order $k$ is an "unordered" selection of $k$ of these.

Common notation: $P(n, k)$ or $P^n_k = n(n-1) \cdots (n-k+1) = n^k$

$$C(n, k) = C^n_k = \binom{n}{k} = \frac{n^k}{k!}.$$  

Example: Given the 5 letters a, b, c, d, e how many ways can we list 3 of the 5 when order is important?

Answer: $P^5_3 = 5^3 = 5 \times 4 \times 3 = 60$.  

Note that each choice of 3 letters (such as a, c, e) results in 6 different results: ace, aec, cae, cea, eac, eca...

Example: Given the 5 letters above how many ways can we choose 3 of the 5 when order is NOT important?

Answer: $\binom{5}{3} = \frac{5^3}{3!} = \frac{5 \times 4 \times 3}{3 \times 2 \times 1} = 10$. In this case we have the 60 that result when we care about order divided by 6 (the number of orderings of 3 fixed letters).
**Definition**

\[ n^k = n \times (n-1) \times \cdots \times (n-k+1) \] for any positive integer \( n \) and for integers \( k \) such that \( 1 \leq k \leq n \). This symbol is pronounced "\( n \) to the \( k \) falling."

Examples: \( 6^3 = 6 \times 5 \times 4 = 120 \).

\( 3^6 \) is not defined (for our purposes).

\( 5^5 = 5! \)

Again: \( P^n_k = n^k \) and \( \binom{n}{k} = \frac{n^k}{k!} \).
Permutations of Multiple Types

The number of permutations of \( n = n_1 + n_2 + \cdots + n_r \) objects of which \( n_1 \) are of one type, \( n_2 \) are of a second type, ..., and \( n_r \) are of an \( rth \) type is

\[
\frac{n!}{n_1!n_2!\cdots n_r!}.
\]

Example: Suppose we have 2 red buttons, 3 white buttons, and 4 blue buttons. How many different orderings (permutations) are there?

Answer: \( \frac{9!}{2!3!4!} = 1260 \).
There are 12 marbles in an urn. 8 are white and 4 are red. The white marbles are numbered w1, w2, ..., w8 and the red ones are numbered r1, r2, r3, r4.

For (a) - (d): Without looking into the urn you draw out 5 marbles.

(a) How many unique choices can you get if order matters? \(12^5 = 95,040\)

(b) How many unique choices can you get if order does not matter? \(\binom{12}{5} = 792\)

(c) How many ways can you choose 3 white marbles and 2 red marbles if order matters? You will fill 5 "slots" by drawing. First determine which two slots (positions) will be occupied by 2 red marbles: \(\binom{5}{2} = 10\). Next multiply by orderings of 3 white and 2 red: \(10 \cdot 8^3 \cdot 4^2 = 40,320\).

(d) How many ways can you choose 3 white marbles and 2 red marbles if order does not matter? \(\binom{8}{3} \cdot \binom{4}{2} = 336\)

(e) How many marbles must you draw to be sure of getting two red ones? 10
Complex Combinations

How many ways are there to create a “full house” (3-of-a-kind plus a pair) using a standard deck of 52 playing cards?

\[
\begin{pmatrix} 13 \\ 1 \end{pmatrix} \begin{pmatrix} 4 \\ 3 \end{pmatrix} \begin{pmatrix} 12 \\ 1 \end{pmatrix} \begin{pmatrix} 4 \\ 2 \end{pmatrix} = 13 \cdot 4 \cdot 12 \cdot 6 = 3,744.
\]

(choose denomination)x(choose 3 of 4 of given denomination)x(choose one of the remaining denominations)x(choose 2 of 4 of this second denomination).

This follows from the multiplication principle (Theorem 2.3.1 in text).
Try these…

Suppose \( \binom{n}{11} = \binom{n}{7} \). What is \( n \)?

Suppose \( \binom{18}{r} = \binom{18}{r - 2} \). What is \( r \)?
Examples*

Consider a machining operation in which a piece of sheet metal needs two identical diameter holes drilled and two identical size notches cut. We denote a drilling operation as \( d \) and a notching operation as \( n \). In determining a schedule for a machine shop, we might be interested in the number of different possible sequences of the four operations. The number of possible sequences for two drilling operations and two notching operations is

\[
\frac{4!}{2!2!} = 6
\]

The six sequences are easily summarized: \( ddnn, dndn, dnnd, nddd, ndnd, nndd \).

Example*

A printed circuit board has eight different locations in which a component can be placed. If five identical components are to be placed on the board, how many different designs are possible?

Each design is a subset of the eight locations that are to contain the components. The number of possible designs is, therefore,

\[
\binom{8}{5} = \binom{8}{3} = \frac{8^3}{3!} = \frac{8 \times 7 \times 6}{3 \times 2 \times 1} = 56.
\]

Sample Space

- **Definition:** The totality of the possible outcomes of a random experiment is called the Sample Space, $\Omega$.
  - **Finite**
    Outcome from one roll of one die $\Rightarrow \Omega = \{1, 2, 3, 4, 5, 6\}$.
  - **Countable**
    The number of attempts until a message is transmitted successfully when the probability of success on any one attempt is $p$
    $\Rightarrow \Omega = \{1, 2, 3, 4, 5, 6, ...\} = \mathbb{Z}^+.$
  - **Continuous (We begin with the discrete cases.)**
    The time (in seconds) until a lightbulb burns out $\Rightarrow \Omega = \{t \in \mathbb{R} : t \geq 0\}$, where $\mathbb{R}$ is the set of all real numbers.
Events

- Definition: An event is a collection of points from the sample space. Example: the result of one throw of die is odd.

- We use sets to describe events. From the die example let the set of "even" outcomes be $E = \{2, 4, 6\}$.

  Let the set of "odd" outcomes be $O = \{1, 3, 5\}$.

- If $\Omega$ is finite or countable, then a “simple” event is an event that contains only one point from the sample space.

  For the die example the simple events are $S_1 = \{1\}, S_2 = \{2\}, ..., S_6 = \{6\}$.

- Suppose we toss a coin until first Head appears. What are the simple events?

- Unless stated otherwise, ALL SUBSETS of a sample space are included as possible events. (Generally we will not be interested in most of these, and many events will have probability zero.)
Describe the sample space and events

- Each of 3 machine parts is classified as either above or below spec.
  - At least one part is below spec.

- An order for an automobile can specify either an automatic or standard transmission, premium or standard stereo, V6 or V8 engine, leather or cloth interior, and colors: red, blue, black, green, white.
  - Orders have premium stereo, leather interior, and a V8 engine.
Describe: sample space and events

- The number of hours of normal use of a lightbulb.
  - Lightbulbs that last between 1500 and 1800 hours.

- The individual weights of automobiles crossing a bridge measured in tons to nearest hundredth of a ton.
  - Autos crossing that weigh more than 3,000 pounds.

- A message is transmitted repeatedly until transmission is successful.
  - Those messages transmitted 3 or fewer times.
Operations on Events

Because the sample space is a set, \( \Omega \), and any event is a subset \( A \subset \Omega \), we form new events from existing events by using the usual set theory operations.

\[ A \cap B \Rightarrow \text{Both } A \text{ and } B \text{ occur.} \]
\[ A \cup B \Rightarrow \text{At least one of } A \text{ or } B \text{ occurs.} \]
\[ \overline{A} \Rightarrow A \text{ does not occur.} \]
\[ S \cap \overline{A} = S - A \Rightarrow S \text{ occurs and } A \text{ does not occur.} \]
\[ \emptyset \Rightarrow \text{the empty set (a set that contains no elements).} \]
\[ A \cap B = \emptyset \Rightarrow A \text{ and } B \text{ are "mutually exclusive."} \]
\[ A \subset B \Rightarrow \text{Every element of } A \text{ is an element of } B, \text{ or, if } A \text{ occurs, } B \text{ occurs.} \]

Review Venn diagrams (in text).
Example

Four bits are transmitted over a digital communications channel. Each bit is either distorted or received without distortion. Let $A_i$ denote the event that the $i$th bit is distorted, $i = 1, 2, 3, 4.$

(a) Describe the sample space.

(b) What is the event $A_1$?

(c) What is the event $A_1 \cup A_2$?

(d) What is the event $A_1'$?
Venn Diagrams
Identify the following events:

(a) $A'$
(b) $A \cap B$
(c) $(A \cap B) \cup C$
(d) $(B \cup C)'$
(e) $(A \cap B)' \cup C$
Mutually Exclusive & Collectively Exhaustive

- A collection of events $A_1, A_2, \ldots$ is said to be mutually exclusive if
  $$A_i \cap A_j = \begin{cases} \emptyset & \text{if } i \neq j \\ A_i = A_j & \text{if } i = j. \end{cases}$$

- A collection of events is collectively exhaustive if
  $$\bigcup_i A_i = \Omega.$$

- A collection of events forms a partition of $\Omega$ if they are mutually exclusive and collectively exhaustive. A collection of mutually exclusive events forms a partition of an event $E$ if
  $$\bigcup_i A_i = E.$$
The sets $A_i$ are "events." No two of them intersect (mutually exclusive) and their union covers the entire sample space.
Probability measure

- We use a probability measure to represent the relative likelihood that a random event will occur.

- The probability of an event $A$ is denoted $P(A)$.

Axioms:

A1. For every event $A$, $P(A) \geq 0$.

A2. $P(\Omega) = 1$.

A3. If $A$ and $B$ are mutually exclusive, then

$$P(A \cup B) = P(A) + P(B).$$

A4. If the events $A_1, A_2, \ldots$ are mutually exclusive, then

$$P \left[ \bigcup_{n=1}^{\infty} A_n \right] = \sum_{n=1}^{\infty} P(A_n).$$
Theorem:

Given a sample space, \( \Omega \), a "well-defined" collection of events, \( \mathbb{F} \), and a probability measure, \( P \), defined on these events then the following hold:

(a) \( P(\emptyset) = 0 \).

(b) \( P[A] = 1 - P[\overline{A}], \forall A \in \mathbb{F} \).

(c) \( P[A \cup B] = P[A] + P[B] - P[A \cap B], \forall A, B \in \mathbb{F} \).

(d) \( A \subseteq B \Rightarrow P[A] \leq P[B], \forall A, B \in \mathbb{F} \).

You must "know" these and be able to use them to solve problems. Don’t worry about proving them.
Applying the Theorem

We roll 1 die and obtain one of the numbers 1 through 6 with equal probability.

(a) What is the probability that we obtain a 7?

The event we want is $\emptyset$; thus, the probability is 0.

(b) What is the probability that we do NOT get a 1?

The event we want is $\{1\}'$ or $\Omega \sim \{1\}$, and $P\left[\{1\}'\right] = 1 - P[\{1\}] = \frac{5}{6}$.

(c) What is the probability that we get a 1 or a 3?

$$P\left[\{1\} \cup \{3\}\right] = P[\{1\}] + P[\{3\}] = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}.$$ 

(d) If $E = \{1, 4, 5, 6\}$, and $E \subseteq G$, what might the event $G$ be?

$G = \{1, 2, 4, 5, 6\}$, $G = \{1, 3, 4, 5, 6\}$, $G = \{1, 2, 3, 4, 5, 6\}$, or $G = E$.

Note that in all of these cases $P[E] \leq P[G]$. 

Applied Statistics: Probability 1-47
Assigning Discrete Probabilities

When there are exactly \( n \) possible outcomes of an experiment, \( x_1, x_2, \ldots, x_n \), then the assigned probabilities, \( p(x_i), i = 1, 2, \ldots, n \), must satisfy the following:

1. \( 0 \leq p(x_i) \leq 1, i = 1, 2, \ldots, n \).
2. \( \sum_{i=1}^{n} p(x_i) = 1 \).

If all of the outcomes have equal probability, then each \( p(x_i) = \frac{1}{n} \); thus, the probability of any particular outcome on the roll of a fair die is \( \frac{1}{6} \).

Suppose, however, we have a biased die and the probability of a 4 is 3 times more likely than the probability of any other outcome. This implies that \( p(1) = p(2) = p(3) = p(5) = p(6) = a \) (for example) and \( p(4) = 3a \).

It follows that \( 8a = 1 \Rightarrow a = \frac{1}{8} \). Thus, \( p(i) = \frac{1}{8}, i \neq 4 \), and \( p(4) = \frac{3}{8} \).
Exercises:

Suppose Bigg Fakir claims that by clairvoyance he can tell the numbers of four cards numbered 1 to 4 that are laid face down on a table. (He must choose all 4 cards before turning any over.) If he has no special powers and guesses at random, calculate the following:

(a) the probability that Bigg gets at least one right
(b) the probability he gets two right
(c) the probability Bigg gets them all right.

Note: assume that Bigg must guess the value of each card before looking at any of the cards.

Hint: What is the sample space? How are the probabilities assigned?
The sample space:

1234 2134 3124 4123  Let $R =$ Number of cards Bigg gets right.
1243 2143 3142 4132  (a) We seek $P[R \geq 1]$.
1324 2314 3214 4213  Suppose that Bigg chooses 1234 (it does not matter).
1342 2341 3241 4231  There is at least one match in the entire first column.
1423 2413 3412 4312  There are three matches in each of the remaining columns.
1432 2431 3421 4321  Thus, there are 15 with at least one match and $P[R] = \frac{15}{24} = \frac{5}{8}$.

Try the other problems.
Complex Combinations

How many ways are there to create a “full house” (3-of-a-kind plus a pair) using a standard deck of 52 playing cards?

\[
\binom{13}{1}\binom{4}{3}\binom{12}{1}\binom{4}{2} = 13 \cdot 4 \cdot 12 \cdot 6 = 3,744.
\]

(choose denomination) \times (choose 3 of 4 of given denomination) \times (choose one of the remaining denominations) \times (choose 2 of 4 of this second denomination).

This follows from the multiplication principle (Theorem 2.3.1 in text).
What is the probability of a “full house”? 

In discrete problems we interpret probability as a ratio:

\[
\frac{\text{number of successful outcomes}}{\text{total number of outcomes}} = \frac{\text{successes}}{\text{successes} + \text{failures}}.
\]

In this case the number of successful outcomes is the number of ways to get a full house (3,744). The total number of outcomes is:

\[
\binom{52}{5} = \frac{52!}{5!} = \frac{52 \cdot 51 \cdot 50 \cdot 49 \cdot 48}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 2,598,960.
\]

Thus, the probability of getting a full house is

\[
\frac{3,744}{2,598,960} = 0.001448.
\]

This is an example of a hypergeometric distribution; we'll study this soon.

A full house happens about once in every 694 hands! This is why people invented wild cards.
Conditional Probability

- The conditional probability of $A$ given that $B$ has occurred is
  \[ P(A|B) = \frac{P(A \cap B)}{P(B)}, \text{ provided } P(B) \neq 0. \]

- **Q1:** What is the probability of obtaining a total of 8 when rolling two dice?

- **Q2:** Suppose you roll two dice that you cannot see. Someone tells you that the sum is greater than 6. What is the probability that the sum is 8?
Dice Problem

Let $A$ be the event of getting 8 on the roll of two dice. Let $B$ be the event that the sum of the two dice is greater than 6. The first question is Find $P(A)$. Here is the sample space:

\begin{align*}
(1,1) & \quad (1,2) & \quad (1,3) & \quad (1,4) & \quad (1,5) & \quad (1,6) \\
(2,1) & \quad (2,2) & \quad (2,3) & \quad (2,4) & \quad (2,5) & \quad (2,6) \\
(3,1) & \quad (3,2) & \quad (3,3) & \quad (3,4) & \quad (3,5) & \quad (3,6) \\
(4,1) & \quad (4,2) & \quad (4,3) & \quad (4,4) & \quad (4,5) & \quad (4,6) \\
(5,1) & \quad (5,2) & \quad (5,3) & \quad (5,4) & \quad (5,5) & \quad (5,6) \\
(6,1) & \quad (6,2) & \quad (6,3) & \quad (6,4) & \quad (6,5) & \quad (6,6)
\end{align*}

Sum=8.

$P(A) = \frac{5}{36}$.

because we are assuming that each outcome pair has the same probability, $1/36$. 

Applied Statistics: Probability 1-54
Dice Problem (conditional)

Here we roll the dice and learn that the sum is greater than 6. Let $B$ represent the event that the sum is greater than 6. With this knowledge the sample space becomes the following:

$$(1,6) \quad (2,5)(2,6) \quad (3,4)(3,5)(3,6) \quad (4,3)(4,4)(4,5)(4,6) \quad (5,2)(5,3)(5,4)(5,5)(5,6) \quad (6,1)(6,2)(6,3)(6,4)(6,5)(6,6)$$

It follows that $P(A \mid B) = \frac{5}{21}$.

Alternatively, by definition of conditional probability, we have

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)}$$

because $A \cap B = A$.

So...the definition makes sense!
Try this.

A university has 600 freshmen, 500 sophomores, and 400 juniors. 80 of the freshmen, 60 of the sophomores, and 50 of the juniors are Computer Science majors. For this problem assume there are NO seniors.

What is the probability that a student, selected at random, is a freshman or a CS major (or both)?

If a student is a CS major, what is the probability he/she is a sophomore?
Use these steps to solve previous slide

1. What is the sample space?
2. What are the events (subsets) of interest?
3. What are the probabilities of the events of interest?
4. What is the answer to the problem?
Curtains...

Suppose you are shown three curtains and told that a treasure chest is behind one of the curtains. It is equally likely to be behind curtain 1, curtain 2, or curtain 3; thus, the initial probabilities are 1/3, 1/3, 1/3 for the treasure being behind each curtain.

Now suppose that Monte Hall, who knows where the treasure is, shows you that the treasure is not behind curtain number 2. The probabilities become $\frac{1}{2}$, 0, $\frac{1}{2}$ right? If you were asked to choose a curtain at this point you would pick either curtain and hope for the best.

Initially we had $P(1) = P(2) = P(3) = \frac{1}{3}$, where $P(i) =$ probability that the prize is behind curtain $i$.

After Monte shows us curtain 2, we have $P(1|\bar{2}) = \frac{P(1 \cap \bar{2})}{P(2)} = \frac{P(1)}{P(2)}$.

$$= \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{1}{2} = P(3|\bar{2}).$$

Note: using the conditional probabilities you do not have to enumerate the conditional sample space.

The formula works again!
Curtains Revisited

We change the game as follows. Again the treasure is hidden behind one of three curtains. At the beginning of the game you pick one of the curtains - say 2. Then Monte shows you that the treasure is NOT behind curtain 3. Now he offers you a chance to switch your choice to curtain 1 or to stay with your original choice of 2. Which should you do? Does the choice make a difference?
Here’s the deal...

After you make your choice (but have not seen what's behind your curtain), the probabilities remain $\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$. Right? This means that the probability that the treasure is behind your curtain is $\frac{1}{3}$ and the probability that it is behind one of the other curtains is $\frac{2}{3}$. Monte will NEVER show you the treasure; thus, even after he shows you one of the curtains, the probability that the treasure is behind the curtain you did not choose is $\frac{2}{3}$. It is twice as likely to be behind the other curtain so... you should always switch!

Moral of the story: be REALLY careful about underlying assumptions about the sample space and how it changes to create the conditional sample space. You are generally assuming the changes are random. Maybe they are not.
Alternate Form

We have seen that the conditional probability of event $A$ given that event $B$ has occurred is: 

$$P(A | B) = \frac{P(A \cap B)}{P(B)}, \text{ provided } P(B) \neq 0.$$ 

Clearly this implies that $P(A \cap B) = P(A | B)P(B)$. This is referred to as the "multiplication rule," and holds even when $P(B) = 0$. Notice that we could also write $P(A \cap B) = P(B | A)P(A)$. Both these equations always hold for any two events. But there is a special case where the conditional probabilities above are not needed.

Note: memorize these conditional probability equations TODAY. They are extremely important.
Independent Events

- Two events $A$ and $B$ are independent iff the probability $P(A \cap B) = P(A)P(B)$.

Example (dice)

- Q1: If one die is rolled twice, is the probability of getting a 3 on the first roll independent of the probability of getting a 3 on the second roll?
- Q2: If one die is rolled twice, is the probability that their sum is greater than 5 independent of the probability that the first roll produces a 1?
** Dice sample spaces (Q1)**

The sample space associated with one roll of a die: \( \Omega_1 = 1, 2, 3, 4, 5, 6 \).

Unless otherwise stated we assume the die is fair so that the probability of any one of the simple events is \( \frac{1}{6} \). The sample space associated with two rolls of one die (or with one roll of a pair of dice):

\[
(1,1)(1,2)(1,3)(1,4)(1,5)(1,6) \\
(2,1)(2,2)(2,3)(2,4)(2,5)(2,6) \\
(3,1)(3,2)(3,3)(3,4)(3,5)(3,6) \\
(4,1)(4,2)(4,3)(4,4)(4,5)(4,6) \\
(5,1)(5,2)(5,3)(5,4)(5,5)(5,6) \\
(6,1)(6,2)(6,3)(6,4)(6,5)(6,6)
\]

Clearly \( P(3, 3) = \frac{1}{36} \). \( P(3 \text{ on first roll}) = P(3 \text{ on second roll}) = \frac{1}{6} \).

Because \( \frac{1}{6} \times \frac{1}{6} = \frac{1}{36} \), the two events are independent.
The probability of getting a 1 on the first die is \( \frac{1}{6} \). Let \( G5 \) be the event that the sum of the two dice is greater than 5 and \( F1 \) be the event that the first roll produces a 1. The sample space is:

\[
(1,1) \ (1,2) \ (1,3) \ (1,4) \ (1,5) \ (1,6) \\
(2,1) \ (2,2) \ (2,3) \ (2,4) \ (2,5) \ (2,6) \\
(3,1) \ (3,2) \ (3,3) \ (3,4) \ (3,5) \ (3,6) \\
(4,1) \ (4,2) \ (4,3) \ (4,4) \ (4,5) \ (4,6) \\
(5,1) \ (5,2) \ (5,3) \ (5,4) \ (5,5) \ (5,6) \\
(6,1) \ (6,2) \ (6,3) \ (6,4) \ (6,5) \ (6,6)
\]

\[
P[F1 \cap G5] = \frac{2}{36} = \frac{1}{18}.
\]

\[
P[G5] = \frac{26}{36} = \frac{13}{18}.
\]

\[
P[F1] = \frac{1}{6}.
\]

Thus, these two events are NOT independent.
Practice Quiz 1 - Explain your work as you have been taught in class.

1. A university has 600 freshmen, 500 sophomores, and 400 juniors. 80 of the freshmen, 60 of the sophomores, and 50 of the juniors are Computer Science majors. For this problem assume there are NO seniors. If a student is a CS major, what is the probability that he/she is a Junior?

2. Evaluate \[ \sum_{i=3}^{\infty} \left( \frac{1}{4} \right)^i. \]

3. What is the probability of drawing 2 pairs in a draw of 5 cards from a standard deck of 52 cards? (A pair is two cards of the same denomination - such as two aces, two sixes, or two kings.)
**Multiplication and Total Probability Rules**

**Multiplication Rule**

\[ P(A \cap B) = P(B|A)P(A) = P(A|B)P(B) \]  

*This slide from *Applied Statistics and Probability for Engineers, 3rd Ed.*, by Douglas C. Montgomery and George C. Runger, John Wiley & Sons, Inc. 2006*
The probability that an automobile battery subject to high engine compartment temperature suffers low charging current is 0.7. The probability that a battery is subject to high engine compartment temperature is 0.05.

Let \( C \) denote the event that a battery suffers low charging current, and let \( T \) denote the event that a battery is subject to high engine compartment temperature. The probability that a battery is subject to low charging current and high engine compartment temperature is

\[
P(C \cap T) = P(C \mid T)P(T) = 0.7 \times 0.05 = 0.035
\]
Multiplication and Total Probability Rules*

Total Probability Rule

Partitioning an event into two mutually exclusive subsets.

Total Probability Rule

Partitioning an event into several mutually exclusive subsets.

*This slide from Applied Statistics and Probability for Engineers, 3rd Ed., by Douglas C. Montgomery and George C. Runger, John Wiley & Sons, Inc. 2006
Problem 2-97a

- A batch of 25 injection-molded parts contains 5 that have suffered excessive shrinkage. If two parts are selected at random, and without replacement, what is the probability that the second part selected is one with excessive shrinkage?

  - $S=\{\text{pairs (f,s) of first-selected, second-selected taken from 25 total with 5 defects}\}$
  - $SD=\{\text{second selected (no replace) is a defect}\}$
  - $FD=\{\text{first selected is a defect}\}$
  - $FN=\{\text{first selected is not a defect}\}$. 
Problem Solution

- We seek \( P[SD] = P[SD \cap FD] + P[SD \cap FN] \)
- This becomes \( P[SD | FN]P[FN] + P[SD | FD]P[FD] \)

\[
= \frac{5}{24} \times \frac{4}{5} + \frac{4}{24} \times \frac{5}{25} = \frac{1}{6} + \frac{1}{30} = \frac{1}{5} = 0.2.
\]
Multiplication and Total Probability Rules*

Total Probability Rule (multiple events)

Assume $E_1, E_2, \ldots, E_k$ are $k$ mutually exclusive and exhaustive sets. Then

$$P(B) = P(B \cap E_1) + P(B \cap E_2) + \cdots + P(B \cap E_k)$$

$$= P(B|E_1)P(E_1) + P(B|E_2)P(E_2) + \cdots + P(B|E_k)P(E_k) \quad (2-8)$$

*This slide from Applied Statistics and Probability for Engineers, 3rd Ed, by Douglas C. Montgomery and George C. Runger, John Wiley & Sons, Inc. 2006*
Total Probability Example

A semiconductor manufacturer has the following data regarding the effect of contaminants on the probability that chips fail.

<table>
<thead>
<tr>
<th>Probability of Failure</th>
<th>Level of Contamination</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>High</td>
</tr>
<tr>
<td>0.01</td>
<td>Medium</td>
</tr>
<tr>
<td>0.001</td>
<td>Low</td>
</tr>
</tbody>
</table>

In a particular production run 20% of the chips have high-level, 30% have medium-level, and 50% have low-level contamination. What is the probability that one of the resulting chips fails?
Bayes Rule

Suppose the events \( B_1, B_2, ..., B_n \) form a partition of the event \( A \), then it follows that

\[
P(B_j \mid A) = \frac{P(A \mid B_j)P(B_j)}{\sum_i P(A \mid B_i)P(B_i)},
\]

provided \( P(A) \neq 0 \).
Exercise:

Moon Systems, a manufacturer of scientific workstations, Produces its Model 13 System at sites A, B, and C, with 20%, 35%, and 45% produced at A, B, and C, respectively. The probability that a Model 13 System will be found Defective upon receipt by a customer is 0.01 if shipped From site A, 0.06 if from B, and 0.03 if from C.

(a) What is the probability that a Model 13 System selected at random at a customer location will be defective?

(b) If a customer finds a Model 13 to be defective, what is the probability that it was manufactured at site B?
Solution (a)

Let $D$ be the event that a Model 13 is found to be defective at a customer site. We want $P[D] = P[D \mid A] P[A] + P[D \mid B] P[B] + P[D \mid C] P[C]$, where $A$ is the event that the Model 13 was shipped from plant $A$ and the events $B$ and $C$ are defined analogously. This is a very important example of conditioning an event $D$ on three events that partition $D$. An equation like this can always be written when a number of events partition an event in which we're interested.

(The general form would be $P[D] = \sum_{i=1}^{n} P[D \mid A_i] P[A_i]$ where $\{A_i\}_i$ form a partition of the event $D$.) Substituting the given numbers we have:

$$P[D] = (0.01)(0.2) + (0.06)(0.35) + (0.03)(0.45) = 0.037,$$

which answers (a).
Solution (b)

(b) If a customer finds a Model 13 to be defective, what is the probability that it was manufactured at site B?

Now we seek \( P(B \mid D) = \frac{P(D \mid B)P(B)}{P(D \mid A)P(A) + P(D \mid B)P(B) + P(D \mid C)P(C)} \)

by Bayes Law. Substituting we get

\[
P(B \mid D) = \frac{(0.06) \times (0.35)}{(0.01) \times (0.2) + (0.06) \times (0.35) + (0.03) \times (0.45)}
= 0.575
\]
Bernoulli Trials

- “Consider an experiment that has two possible outcomes, success and failure. Let the probability of success be $p$ and the probability of failure be $q$ where $p+q=1$. Now consider the compound experiment consisting of a sequence of $n$ independent repetitions of this experiment. Such a sequence is known as a sequence of Bernoulli Trials.”

- The probability of obtaining exactly $k$ successes in a sequence of $n$ Bernoulli trials is the binomial probability

$$p(k) = \binom{n}{k} p^k q^{n-k}.$$ 

- Note that the sum of the probabilities $\sum_k p(k) = 1$. Thus they are said to form a probability distribution.
Probable

Probability Distribution

When we take a discrete or countable sample space $\Omega = \{s_1, s_2, \ldots\}$ and assign probabilities to each of the possible simple events: $P(\{s_1\}) = p_1, P(\{s_2\}) = p_2, \ldots$, we have created a probability distribution. (Think that you have "distributed" all of the probability over all possible events.) As an example, if I toss a coin one time then $P(H) = \frac{1}{2}$ and $P(T) = \frac{1}{2}$ represents a probability distribution.

The single coin toss distribution also is an example of a Bernoulli trial because it has only two possible outcomes (generally called "success" or "failure).
The binomial probabilities are defined by \( p(k) = \binom{n}{k} p^k q^{n-k} \), where \( p \) is the probability of success and \( q \) is the probability of failure in \( n \) Bernoulli trials. Suppose we toss a coin 10 times and we want the total number of heads. Then \( p = P[H], q = P[T], n = 10 \). Using the above formula we obtain the probabilities:

<table>
<thead>
<tr>
<th>( k )</th>
<th>( p(k) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0000</td>
</tr>
<tr>
<td>1</td>
<td>0.009766</td>
</tr>
<tr>
<td>2</td>
<td>0.0440</td>
</tr>
<tr>
<td>3</td>
<td>0.1167</td>
</tr>
<tr>
<td>4</td>
<td>0.205</td>
</tr>
<tr>
<td>5</td>
<td>0.246</td>
</tr>
<tr>
<td>6</td>
<td>0.205</td>
</tr>
<tr>
<td>7</td>
<td>0.1167</td>
</tr>
<tr>
<td>8</td>
<td>0.0440</td>
</tr>
<tr>
<td>9</td>
<td>0.009766</td>
</tr>
<tr>
<td>10</td>
<td>0.0000</td>
</tr>
</tbody>
</table>
Regarding Parameters

Notice that the binomial distribution is completely defined by the formula for its probabilities, \( p(k) = \binom{n}{k} p^k q^{n-k} \), and by it "parameters" \( p \) and \( n \).

The binomial probability equation never changes so we regard a binomial distribution as being defined by its parameters. This is typical of all probability distributions (using their own parameters, of course).

One of the problems we often face in statistics is estimating the parameters after collecting data that we know (or believe) comes from a particular probability distribution (such as the \( p \) and \( n \) for the binomial). Alternatively, we may choose to estimate "statistics" such as mean and variance that are functions of these parameters. We'll get to this, after we consider random variables and the continuous sample space.
Example (from Trivedi*)

Consider a binary communication channel transmitting coded words of $n$ bits each. Assume that the probability of successful transmission of a single bit is $p$ and that the probability of an error is $q = 1 - p$. Assume also that the code is capable of correcting up to $e$ errors, where $e \geq 0$. If we assume that the transmission of successive bits is independent, then the probability of successful word transmission is:

$$P_w = P\{e \text{ or fewer errors in } n \text{ trials}\}$$

$$= \sum_{i=0}^{e} \binom{n}{i} q^i p^{n-i}.$$  

Notice that a "success for the Binomial distribution" means getting an error, which has probability $q$.

Example

A communications network is being shared by 100 workstations. Time is divided into intervals that are 100 ms long. One and only one workstation may transmit during one of these time intervals. When a workstation is ready to transmit, it will wait until the beginning of the next 100 ms time interval before attempting to transmit. If more than one workstation is ready at that moment, a collision occurs; and each of the \( k \) ready workstations waits a random amount of time before trying again. If \( k = 1 \), then transmission is successful. Suppose the probability of a workstation being ready to transmit is \( p \). Show how probability of collision varies as \( p \) varies between 0 and 0.1.

\[
pColl(p) := 1 - \text{dbinom}(0, 100, p) - \text{dbinom}(1, 100, p)
\]
Practice Quiz 2

A partial deck of playing cards (fewer than 52 cards) contains some spades, hearts, diamonds, and clubs (NOT 13 of each "suit"). If a card is drawn at random, then the probability that it is a spade is 0.2. We write this as P[Spade]=0.2. Similarly, P[Heart]=0.3, P[Diamond]=0.25, P[Club]=0.25. Each of the 4 suits has some number of “face” cards (King, Queen, Jack). If the drawn card is a spade, the probability is 0.25 that it is a face card. If it is a heart, the probability is 0.25 that it is a face card. If it is a diamond, the probability is 0.2 that it is a face card. If it is a club, the probability is 0.1, that it is a face card.

1. What is the probability that the randomly drawn card is a face card?
2. What is the probability that the card is a Heart and a face card?
3. If the card is a face card, what is the probability that it is a spade?