

WEEK 2: CH 2, PROBABILITY

TABLE OF CONTENTS

	Page:
List of Figures.	1
List of Tables.	1
List of m.files.	2
Assignment.	2
Required reading.	2
Understanding by Design Templates.	3
Understanding By Design Stage 1 — Desired Results Week 3	3
Understanding by Design Stage 2 — Assessment Evidence Week 2.	3
Introduction.	5
Bayes Theorem & MCAS failure.	5
What is the probability of being classified an MCAS Failure if you are not a failure: Bayes Theorem.	6
Annotated outline (with Matlab scripts) for Larsen & Marx Chapter 2.	7
References.	44
Index.	44

List of Figures

Figure 1. Figure 2.2.2 From Larsen & Marx (2006, p 29)	12
Figure 2. Figure 2.2.3 From Larsen & Marx (2006, p 30)	14
Figure 3. Figure 2.2.4 From Larsen & Marx (2006, p 33)	15
Figure 4. Figure 2.2.7 From Larsen & Marx (2006, p 33)	15
Figure 5. Figure 2.3.2 From Larsen & Marx (2006, p 39)	18
Figure 6. Example 2.3.4 From Larsen & Marx (2006, p 40)	19
Figure 7. Figure 2.6.14 from Larsen & Marx (2006, p 104)	39

List of Tables

Table 1. An analysis of the 10 th grade 1998 MCAS math classification analysis.	5
Table 2.4.3.	27
Table 2.4.4.	28
Table 2.8.2 Craps.	33

List of m.files

LMex020201_4th.....	8
LMex020202_4th.....	9
LMex020203_4th.....	10
LMex020205_4th.....	11
LMex020206_4th.....	11
LMex020209_4th.....	12
LMex020209_4th.....	13
LMex020210_4th.....	14
LMex020213_4th.....	15
LMex020301_4th.....	18
LMex020303_4th.....	19
LMex020401_4th.....	21
LMex020402_4th.....	22
LMex020411_4th.....	23
LMex020412_4th.....	25
LMex020413_4th.....	26
LMex020414_4th.....	26
LMex020415_4th.....	27
LMex020503_4th.....	28
LMex020507_4th.....	29
function G=gini(DATA).....	29
LMex020509_4th.....	30
crapstableMOF65.....	31
function [Pexact,Pnumber,Odds]=crapsmof65(number,H).....	31
LMex020606_4th.....	35
LMex020611_4th.....	36
Example 2.6.13 How many ways to arrange 1 nickel, 3 dimes, and 2 quarters?.....	37
function [HB,E]=brillouin(DATA).....	37
Example 2.6.14 Huygens anagram.....	38
LMex020616_4th.....	38
function V=palindrom(N).....	38
LMex020617_4th.....	39
LMex020701_4th.....	41
LMex020705_4th.....	41
LMex020706_4th.....	41
function [P,PM]=birthdays(k,trials).....	42
LMex020707_4th.....	42

Assignment

Required reading

- ! Larsen, R. J. and M. L. Marx. 2006. An introduction to mathematical statistics and its applications, 4th edition. Prentice Hall, Upper Saddle River, NJ. 920 pp.
- " Read Sections 2.1-2.8

Understanding by Design Templates

Understanding By Design Stage 1 — Desired Results Week 3

Larsen & Marx (2006) Chapter 2 Probability

G Established Goals (including those from the MA Dept. of Education MCAS 11th & 12th grade standards, <http://www.doe.mass.edu/frameworks/current.html>)

- Calculate & apply unconditional and conditional probabilities
- Apply uniform, binomial, hypergeometric and multinomial distributions to the solutions of problems (DOE 12.D.4)
- Use combinatorics to solve problems, in particular, to compute the probability of compound events (DOE 12.D.6)
- Compare the results of simulations with predicted probabilities (DOE 12.D.7)

U Understand

- P values from Monte Carlo simulations converge to analytical probabilities
- That there are two non-mutually exclusive schools of statistics: frequentist & Bayesian
- $N! = (N+1)$, so $\text{gammaln}(N+1)$ is the key to solving combinatorial problems
- Many biodiversity indices are based on combinatorics

Q Essential Questions

- In 2-child families, what is the probability that both are boys if at least one is a boy?
- If Monty Hall offers, should you switch doors in ‘The Price is Right’?
- For men, given a positive PSA test, should you have your prostate removed?
- Why isn’t this course being offered using Bayesian methods?
- How many students must be in a class to have better than even odds that 2 or more share a birthday?
- Why didn’t Sanders’ 1968 biodiversity rarefaction formula work?

K Students will know how to define (in words or equations)

- **Bayes Theorem, binomial coefficient, combination, complement, conditional probability, event, experiment, false negative, false positive**, Hurlbert’s $E(S_n)$, **independence, intersection**, Monte Carlo simulation, **multiplication rule, multinomial coefficient**, null set, odds, **permutations**, prevalence rate, **probability**, sample outcome, sample space, sensitivity, specificity, union, **Venn diagram**, yield or predictive value

S Students will be able to

- Plot Venn diagrams in Matlab
- Analyze problems with Bayes’ theorem
- Count outcomes using the multiplication rule, permutations & combinations
- Write combinatorial problems using Matlab’s gammaln function
- Analyze biodiversity using Matlab with Brillouin’s H & Shannon’s H

Understanding by Design Stage 2 — Assessment Evidence Week 2

LM Chapter 2 (21-127) Probability

Post in the discussion section by 6/15/11 W

- Marshall Nirenberg discovered the **genetic code**, refuting the code proposed by George Gamow. Gamow's proposed code was based on the 24 3-nucleotide permutations of AGCU. Is it chance that all life is based on Nirenberg's code and not Gamow's?
- **HW2 Problems due Wednesday 6/15/11 W 10 PM**
 - Each problem must be solved using Matlab
 - **Basic problems (Pick 4 of 5 problems 10 points)**
 - **Problem 2.2.4 P. 26 Outcomes in a deck of cards**

Two cards are dealt from a standard 52-card deck. Let A be the event that the sum of two cards is eight. Write a Matlab program to list the outcomes in A . Use LMex020203_4th.4 and LMex020201_4th.m as paradigms. I used this as my opening line of a 6-line program:
deck=repmat(1:52/4,1,4)
 - **Problem 2.3.8 P.41 Venn diagrams**
 - Use venn.m (from the Mathworks website) or from Gallagher's EEOS601 files
 - Use Example 2.2.12 & 2.3.1 as models
 - **Problem 2.4.18 P. 53**

Use LMex020202_4th.m as a model
 - **Problem 2.4.46 P. 68 Murder She Wrote**

Use Example 2.4.12 (LMex020412_4th.m) as a model (or 2.4.13, 2.4.14)
 - **Problem 2.6.34 P. 105.** Which state can generate more permutations: Tennessee or Florida?
 - **Advanced problems (2.5 points each)**
 - Continue the analysis of Example 2.2.9 on p. 29 and Gallagher's LMex020209_4th.m. Let C be the event that the sum of the faces is even. Analyze $A \cap C$ and $B \cap C$ with a Matlab program that counts and displays the outcomes.
 - Example 2.4.2 p 44. Modify Gallagher's Matlab program LMex020402_4th.m to produce a Monte Carlo simulation of the two-boys problem. Use LMex020303_4th.m as a model.
 - **Master problems (2.5 points each)**
 - **Problem 2.3.10.** Write a Matlab program to solve $A \cup B$ exactly and as a Monte Carlo simulation. Use LMex020209_4th.m and LMex020303_4th.m as model
 - Modify the MCASQ9 code from HO4 to find the area of the outer circle. Do not use pi or any approximations for pi in your solution. Use the Nahin program and Bevington & Taylor figure as a model.

Introduction

Chapter 2 introduces the probability theory that is the foundation of all statistical analysis. The list of terms that must be mastered is daunting (see the list under K in the Understanding by Design template). On top of the conceptual difficulty of mastering probability is the issue of programming these techniques in Matlab. The Matlab programs for the final 8 chapters of the text are simple: enter the data and use 2 or 3 Matlab functions to perform the statistical analysis and graph the results. For chapters 2-3, it becomes very difficult to even conceive how to turn word problems into Matlab algorithms. These programs tend to use indexing to select sample outcomes and determine probabilities. All I can say is that try to work through this section of the class as well as you can. Things will get easier as these concepts and Matlab programming tricks become more familiar.

Bayes Theorem & MCAS failure

The MA Dept of Educations' Technical Appendix C for the 1998 math 10th grade MCAS reveals a disturbing fact. Even if the testing theory underlying MCAS is correct, almost 17% of the students who should have scored in the 'Needs Improvement' category will fail MCAS. About 9% of all the students who should have passed MCAS will receive failing scores.

Table 1. An analysis of the 10th grade 1998 MCAS math classification analysis shown in Appendix C of the October 1999 Technical report for the 1998 MCAS exam. The first table shown under the heading Step 4 shows the data as presented in the technical Appendix C. I converted this table to conditional probabilities, by dividing each probability by the row marginal total. This shows what percentage of students in each true category are misclassified. 16.53% (4.27/25.86) of the students in the "Needs Improvement" category were misclassified as failing.

Step 4		Predicted Classification				
		Fail	NI	Prof	Adv	Marginal
True Status	Fail	46.95%	4.04%	0.01%	0.00%	51.00%
	Needs Improvement	4.27%	17.90%	3.68%	0.01%	25.86%
	Proficient	0.00%	2.96%	12.68%	1.97%	17.61%
	Advanced	0.00%	0.00%	1.08%	4.44%	5.52%
Marginal		51.22%	24.90%	17.45%	6.41%	99.99%
Misclassification rates (bolded in red)						
		Predicted				
		Fail	NI	Prof	Adv	Marginal
True Status	Fail	92.06%	7.91%	0.03%	0.00%	100.00%
	Needs Improvement	16.53%	69.21%	14.23%	0.03%	100.00%
	Proficient	0.01%	16.83%	71.97%	11.19%	100.00%
	Advanced	0.00%	0.03%	19.60%	80.37%	100.00%

What is the probability of being classified an MCAS Failure if you are not a failure: Bayes Theorem

The Problem

Bayes' Theorem

Let $\{ A_1 \}_n^{i=1}$ be a set of n events,
 each with positive probability,
 that partitions S in such a way that

$$\bigcup_{i=1}^n A_i = S$$

and $A_i \cap A_j = \emptyset$ for $i \neq j$.

For any event B (also defined on S),
 where $P(B) > 0$,

$$P(A_j|B) = \frac{P(B|A_j)P(A_j)}{\sum_{i=1}^n P(B|A_i)P(A_i)}$$

for any $i \leq j \leq n$.

What is the probability of receiving a failing grade on the MCAS test if your true abilities are in the passing range? Let Event A be the probability of a student getting a failing score on MCAS and Event B be the probability that the student's true ability is in the passing range. Our goal is to determine $P(A/B)$. We using Bayes theorem.

$$\frac{P(\text{MCAS Failure} | \sim\text{Failure}) = P(\sim\text{Failure}|\text{MCAS Failure}) P(\text{MCAS Failure})}{P(\sim\text{Failure} | \text{MCAS Failure}) P(\text{MCAS Failure}) + P(\sim\text{Failure} | \sim\text{MCAS Failure}) P(\sim\text{MCAS Failure})}$$

From the upper part of Table 1:

Numerator

$$P(B|A)=P(\sim\text{Failure}|\text{MCAS Failure}) = 4.27/51.22= 1- (46.95/51.22) = 0.0834$$

$$P(A)= P(\text{MCAS Failure}) = 0.5122$$

$$P(B|A)*P(A)=0.0834*0.5122 = 0.0427$$

Denominator

Do the calculation from the definition of conditional probability:

$$P(\sim\text{Failure} | \sim\text{MCAS Failure}) = P(\text{Not a Failure}) \cap P(\sim\text{MCAS Failure}) / P(\sim\text{MCAS Failure});$$

$$((.2586-.0427)+(0.1761-0)+(0.0552-0)) / (1-.5122)$$

$$= 0.9168$$

or by summing the individual probabilities:

$$\text{denominator} = P(B|A)*P(A)+P(B|\sim A)*P(\sim A)$$

$$P(B | \sim A) = (0.1790+0.0368+0.0001+0.0296+0.1268+0.0197+0.0108+0.0444) / (1-.5122)$$

$$= 0.9168$$

$$P(\sim A) = P(\sim \text{MCAS Failure}) = 1 - 0.5122 = 0.4878$$

Combine these probabilities in Bayes' theorem

$$P = 0.0834 * 0.5122 / (0.0834 * 0.5122 + 0.9168 * 0.4878) = 0.0427 / 0.4899$$

$$P = 0.0872$$

Therefore, there is about a 9% chance that if a student received his MCAS failure notice for the 1998 exam (<220 on a 200 to 400 scale), that the correct classification should have been "not failing."

This misclassification rate is based on the latest information available, for the 1998 test. Since this is the test on which all other MCAS tests are based, it is highly unlikely that this misclassification rate could be less on later tests. Scaling a later test to resemble the 1998 scores introduces a large additional source of variability, which should increase the misclassification rate. In the Spring 2001 test, the modal distribution of failures moved closer to the 220 cutoff. This would tend to increase the misclassification rate. To do this calculation requires the appendices for the MCAS technical reports. The technical reports have only been published for the 1998 and 1999 exams, and the DOE web site doesn't include the technical appendices for the 1999 MCAS exam.

Annotated outline (with Matlab scripts) for Larsen & Marx Chapter 2

2 Chapter 2: Probability

Pierre de Fermat (1601-1605)

Blaise Pascal (1623-1662)

2.1 INTRODUCTION

2.1.1 Examples

2.1.1.1 NASA Challenger: if the odds of a Challenger type disaster are 1 in 78, what are the odds of getting a disaster in the 20 launches over a 5-year period?

2.1.1.2 Polygraph

2.1.1.3 Baseball player: What are the odds of getting a hit in 30 games?

2.1.1.4 Births

2.1.2 The Evolution and Definition of Probability (p. 22)

2.1.2.1 Cardano's classic definition of probability

2.1.2.1.1 "...the probability of an *event* comprised of m outcomes is the ratio m/n , where n is the total number of equally likely outcomes."

2.1.2.2 Von Mises's empirical probabilities. Repeating an experiment over and over

2.1.2.3 Andrei Kolmogorov in 1933

2.2 SAMPLE SPACES AND THE ALGEBRA OF SETS (p 24)

2.2.1.1 Definitions

- 2.2.1.1.1 “By **experiment** we will mean any procedure that (1) can be repeated, theoretically, an infinite number of times; and (2) has a well-defined set of possible outcomes” (**Larsen & Marx 2006 p. 24**)
- 2.2.1.1.2 **Sample outcome** “Each of the potential eventualities of an **experiment** is referred to as a **sample outcome**, s , and their totality is referred to as a **sample space**, S . To signify the membership of s in S , we write $s \in S$. Any designated collection of **sample outcomes**, including individual outcomes, the entire sample space, and the null set, constitutes an **event**. The latter is said to occur if the outcome of the **experiment** is one of the members of the **event**.” (**Larsen & Marx 2006, p. 24**)

Example 2.2.1 Tossing 3 coins

```
% LMex020201_4th.m
% Example 2.2.1 p 24 in
% Larsen & Marx (2006) Introduction to Mathematical Statistics, 4th edition
% Written by Eugene.Gallagher@umb.edu for EEOS601 1/13/11
% http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html
% 3 Coin tosses of fair coin. List the sample space S and the sample
% outcomes for event A: majority of faces heads.
coin=['H';'T']
% create the sample space
tosses=3
S=repmat(' ',2^tosses,3);
i=1;
for face1=1:2
    for face2=1:2
        for face3=1:2
            S(i,:)= [coin(face1) coin(face2) coin(face3)];
            i=i+1;
        end
    end
end
disp('The sample space, S:')
S
% find sample outcomes in sample space S that make up the event A: Majority
% of coins show heads.
[r,c]=size(S);i=find(sum(S=='H')>c/2);
fprintf('%1.0f of the sample outcomes in Sample Space S constitute the event A:\n',...
```



```
length(i))  
A=S(i,:)
```

Example 2.2.2 Rolling two dice

```
% LMex020202_4th.m  
% Example 2.2.2 Throwing 2 dice p  
% Larsen & Marx (2006) Example 2.2.2, page 24 Figure 2.2.1  
% Writtten by Eugene Gallagher, Eugene.Gallagher@umb.edu  
% http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html  
% Written 9/17/2010 & last revised 1/13/2011  
% What is the probability of rolling a 7 with two dice  
% The dice are colored red and green, but that is irrelevant here.  
green=1:6  
red=green'  
% tile rows down 6 times  
green= repmat(green,6,1) % repmat is a built-in function  
% for earlier versions of matlab, I used rowcopy and colcopy.m  
% tile columns across 6 times  
red= repmat(red,1,6) % Note that the 1 column green is replaced  
% by a 6x6 matrix.  
sumdice=red+green  
[r,c]=size(sumdice) % returns the number of rows and columns  
fprintf('Sample Space, with each sample outcome an ordered pair:\n')  
fprintf('%1.0f, %1.0f) (%1.0f, %1.0f) (%1.0f, %1.0f) (%1.0f, %1.0f) (%1.0f, %1.0f) (%1.0f, %1.0f) (%1.0f, %1.0f)\n',[green(:);red(:)])  
i=find(sumdice==7); % indices for event A: sum of faces=7  
length(i)  
fprintf(...  
'%1.0f of the %2.0f sample outcomes in sample space S constitute the event A, rolling a 7:\n',...  
length(i),r*c)  
fprintf('%1.0f,%1.0f)\n',[red(i) green(i)])  
% Not specified the problem, but here's the probability of rolling a 7:  
P=length(i)/(r*c) % Probability of rolling a 7  
% It is more accurate to express simple probabilities  
% as fractions: 0.1667 is not exactly equal to 1/6  
% Matlab can express results in fractional form  
% by typing format rat, short for rational number format  
format rat  
disp('The probability of rolling a 7 with two dice is:')  
P  
% Here is one of many 1-line solutions to the problem  
red=repmat(1:6,6,1);c=red+red';P=sum(c(:)==7)/length(c(:));  
% typing format returns Matlab to its numerical default  
format  
% Translation of the 1-line program  
% 1st repmat copies the row vector 1:6 down producing 6 rows of 1:6
```

```
% The sample space, c is created by adding the results of the red die to  
% its transpose red'  
% c represents the sample space, produced by summing 1:6 for first die &  
% 1:6 for 2nd die  
% c(:) converts any matrix to a single column of numbers  
% c(:)==7, a logical statement, converts c to a set of 1's and 0's  
%      a 1 is entered if c equals 7, a zero if c(i) not equal to zero  
% sum(c(:)==7) provides the number of elements equal to 9  
% length(c(:)==7) provides the number of events in the sample space, 36.
```

Example 2.2.3 Newscasting positions

```
% LMex020203_4th.m  
% Example 2.2.2 Throwing 2 dice p  
% Larsen & Marx (2006) Example 2.2.2, page 24 Figure 2.2.1  
% Written by Eugene Gallagher, Eugene.Gallagher@umb.edu  
% http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html  
% Written 9/17/2010 & last revised 1/13/2011  
% What is the probability of rolling a 7 with two dice  
% The dice are colored red and green, but that is irrelevant here.  
% The curly brackets make V a cell-structured matrix allowing the  
% labels to be sorted as a unit.  
V=[{'W1'};{'W2'};{'W3'};{'M1'};{'M2'}];  
fprintf('Sample Space, with each sample outcome an ordered pair:\n')  
% This is the built-in Matlab combination function. Since V is an input  
% vector, not a scalar, all of the combinations of size 2 are output.  
nchoosek(V,2)
```

Example 2.2.4 Infinite sample space

```
% LMex020204_4th.m  
% Example 2.2.1 p 24 in  
% Larsen & Marx (2006) Introduction to Mathematical Statistics, 4th edition  
% Written by Eugene.Gallagher@umb.edu for EEOS601 1/13/11  
% http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html  
% A coin is tossed until the first head appears  
% This doesn't directly address the issue posed in the text, but instead  
% shows another type of control loop, capable, in theory of producing an  
% infinite string of coin tosses (all heads)  
coin=['H';'T']  
% create the sample space  
% S=cell(1);  
i=1;  
face='H';  
S='';  
while face=='H'  
    face=coin(randi(2));  
    S(i)=face;  
    i=i+1;
```

```
end  
disp('The sample space, S:')  
S
```

```
Example 2.2.5 Quadratic equation,  
% LMex020205_4th.m  
% Example 2.2.5 page 25-26 in  
% Larsen & Marx (2006) Introduction to Mathematical Statistics, 4th edition  
% A computer programmer is runing a subroutine that solves a general  
% quadratic equation 'a*x^2 + b*x + c' Her experiment consists of choosing  
% values for the three coefficients a, b, and c. Define (1) S and (2) the  
% event A: Equation has two equal roots.  
syms a b c x A u v w z  
A = solve(a*x^2 + b*x + c);  
fprintf('Membership is A is contingent upon a, b and c satisfying:\n')  
a=solve(A(1)-A(2),a)  
b=solve(A(1)-A(2),b)  
c=solve(A(1)-A(2),c)
```

Questions Page 26-28

2.2.1 Graduating engineer
2.2.2 3 dice problem Fall 2010
2.2.3 Chips in urns
2.2.4 52-card deck: outcomes summing to 8 Summer 2011
Women picking suspects in a lineup
Quadratic equation (look up solution form)
Prisoners and urns

2.2.2 Unions, intersections and complements

Definition 2.2.1 Let A and B be any two events defined over the same sample space S . Then
The **intersection** of A and B , written $A \cap B$, is the event whose outcomes belong
to both A and B . The **union** of A and B , written $A \cup B$, is the event whose
outcomes belong to either A or B or both.

```
Example 2.2.6            A single card is drawn from a poker deck. Not an easy Matlab problem.  
% LMex020206_4th.m  
% Example 2.2.7 page 28 in  
% Larsen & Marx (2006) Introduction to Mathematical Statistics, 4th edition  
% Written by Eugene.Gallagher@umb.edu 1/14/11 31  
% http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html  
% A single card is drawn from a poker deck. A is the event that an Ace has  
% been selected and B is the event "Heart is drawn" Find A intersection B  
% and A union B. Let C be the event a club is drawn. Find B U C and B^C.  
deckn= repmat(1:13,1,4);  
deckf= repmat('HCDS',1,13);  
Ai=find(deckn==1);Bi=find(deckf=='H');  
i=intersect(Ai,Bi);  
fprintf('The intersection of A & B has %2.0f outcome(s):\n',length(i))
```

```
fprintf('1c%2.0f\n',[deckf(i)' deckn(i)'])
i=union(Ai,Bi);
fprintf('The union of A & B has %2.0f outcome(s):\n',length(i))
fprintf('1c%2.0f\n',[deckf(i)' deckn(i)'])
Ci=find(deckf=='C');
i=union(Bi,Ci);
fprintf('The union of B & C has %2.0f outcome(s):\n',length(i))
fprintf('1c%2.0f\n',[deckf(i)' deckn(i)'])
i=intersect(Bi,Ci);
fprintf('The intersection of B & C has %2.0f outcome(s):\n',length(i))
fprintf('1c%2.0f\n',[deckf(i)' deckn(i)'])
```

Example 2.2.7 Let A be the set of x 's for which $x^2 + 2x = 8$; let B be the set for which $x^2 + x = 6$. Find $A \cap B$ and $A \cup B$.

```
% LMex020207_4th.m
% Example 2.2.7 page 28 in
% Larsen & Marx (2006) Introduction to Mathematical Statistics, 4th edition
% Written by Eugene.Gallagher@umb.edu 1/14/11 41
% Let A be the set of x's for which  $x^2+2*x=8$ ; let B be the set for which
%  $x^2+x=6$ . Find  $A \cap B$  and  $A \cup B$ . This will be an application of Matlab's
% symbolic math equation solver
syms A B x
A=solve(x^2+2*x-8)
B=solve(x^2+x-6)
AintB=solve(x^2+2*x-8,x^2+x-6)
AunionB=unique([A B])
```

Example 2.2.8 Consider the electrical circuit pictured in Figure 2.2.2. Let A_i denote the event that switch i fails to close, $i=1,2,3,4$. Let A be the event "Circuit is not completed." Express A in terms of the A_i 's.

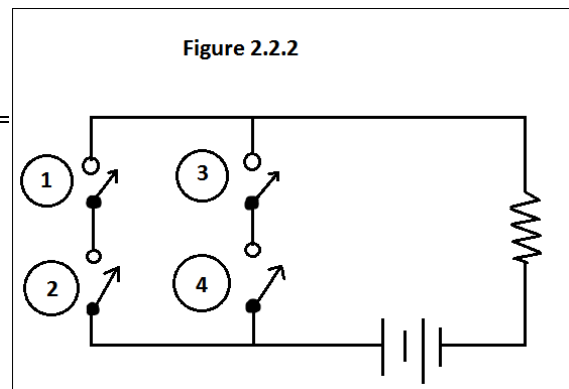


Figure 1. Figure 2.2.2

Call the ① and ② switches, line a; call the ③ and ④ switches, line b; By inspection, the circuit fails only if *both* line a and line b fail. But line a fails only if either ① or ② (or both) fail. That is, the event that line a fails is the union $A_1 \cup A_2$. Similarly the failure of line b is the union $A_3 \cup A_4$. The event that the circuit fails, then, is the intersection:

$$A = (A_1 \cup A_2) \cap (A_3 \cup A_4)$$

[See Nahin problem 3: Will the light bulb glow?]

Example 2.2.9

```
% LMex020209_4th.m
% Example 2.2.9 page 28 in
% Larsen & Marx (2006) Introduction to Mathematical Statistics, 4th edition
```

```
% Written by Eugene.Gallagher@umb.edu 1/14/11 44
% http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html
% Let A be the event that the sum of the faces is odd. Let B be the event
% that the two faces themselves are odd. What is the intersection?
% Use the dice example LMex020202_4th.m as a paradigm
green=1:6
red=green'
% tile rows down 6 times
green= repmat(green,6,1) % repmat is a built-in function
% for earlier versions of matlab, I used rowcopy and colcopy.m
% tile columns across 6 times
red= repmat(red,1,6) % Note that the 1 column green is replaced
% by a 6x6 matrix.
sumdice=red+green
[r,c]=size(sumdice); % returns the number of rows and columns
i=find(mod(sumdice,2)); % Returns the remainder after dividing by two
[A,m]=sort(sumdice(i)); % finds all of the odd sums, at location i
fprintf('Event A consists of the following dice and their sum:\n');
disp(' red green sum')
disp([red(i(m)) green(i(m)) A])
j=find(mod(red,2)& mod(green,2));
B=[red(j) green(j)]
AintB=intersect(i,j);
if isempty(AintB)
    fprintf('The intersection of A and B is empty.\n')
else
    fprintf('The intersection of A and B has %2.0f outcomes:\n',...
        length(AintB))
    disp(' red green sum')
    disp([red(AintB) green(AintB) sumdice(AintB)])
end
```

Definition 2.2.2. Events A and B defined over the same sample space are said to be **mutually exclusive** if they have no outcomes in common — that is, if $A \cap B = \emptyset$, where \emptyset is the null set.

Example 2.2.9 Throwing two dice & analyzing the intersections of sets in Matlab

```
% LMex020209_4th.m
% Example 2.2.9 page 28 in
% Larsen & Marx (2006) Introduction to Mathematical Statistics, 4th edition
% Written by Eugene.Gallagher@umb.edu 1/14/11
% http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html
% Let A be the event that the sum of the faces is odd. Let B be the event
% that the two faces themselves are odd. What is the intersection?
% Use the dice example LMex020202_4th.m as a paradigm
green=1:6
red=green'
% tile rows down 6 times
```

```

green= repmat(green,6,1) % repmat is a built-in function
% for earlier versions of matlab, I used rowcopy and colcopy.m
% tile columns across 6 times
red= repmat(red,1,6) % Note that the 1 column green is replaced
% by a 6x6 matrix.
sumdice=red+green
[r,c]=size(sumdice); % returns the number of rows and columns
i=find(mod(sumdice,2)); % Returns the remainder after dividing by two
[A,m]=sort(sumdice(i)); % finds all of the odd sums, at location i
fprintf('Event A consists of the following dice and their sum:\n');
disp(' red green sum')
disp([red(i(m)) green(i(m)) A])
j=find(mod(red,2)& mod(green,2));
B=[red(j) green(j)]
AintB=intersect(i,j);
if isempty(AintB)
    fprintf('The intersection of A and B is empty.\n')
else
    fprintf('The intersection of A and B has %2.0f outcomes:\n',...
        length(AintB))
    disp(' red green sum')
    disp([red(AintB) green(AintB) sumdice(AintB)])
end
  
```

Definition 2.2.3. Let A be any event defined on a sample space S . The **complement** of A , written A^c is the event consisting of all the outcomes in S other than those contained in A .

Example 2.2.10 Let A be the set of $\{x, y\}$ for which $x^2 + y^2 < 1$. Sketch the region in the x, y plane corresponding to A^c .

```

% LMex020210_4th.m
% Example 2.2.10 page 30 in
% Larsen & Marx (2006) Introduction to
Mathematical Statistics, 4th edition
% Written by Eugene.Gallagher@umb.edu 1/14/11
% http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html
syms x y; ezplot(x^2+y^2-1); axis('square')
% Go into the Matlab editor and fill the area outside the circle to make a
% proper 1-circle venn diagram
  
```

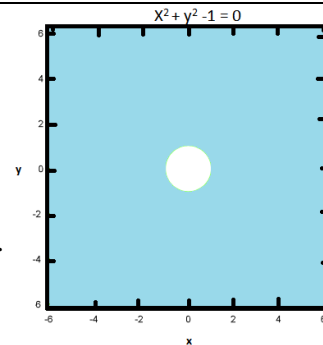


Figure 2.

Example 2.2.11 Not a Matlab problem.

Questions p 30

Sketch the regions in the x - y plane corresponding to $A \cup B$ and $A \cap B$ if ...
 Referring to example 2.2.7

Find

An electronic system has four components divided into two pairs. The two components of each pair are wired in parallel; the two pairs are wired in series. Let A_{ij} denote the event “ith

component in j th pair fail" $i=1,2; j=1,2$. Let A be the event "System fails" Write A in terms of the A_{ij} 's.

2.2.3

2.2.4

2.2.5 Expressing ideas graphically Venn diagrams

Note that there is a nice Matlab program called Venn.m on the User contributed m.files that can be used for these examples

Fig. 2.2.4

Example 2.2.12 Venn diagrams

```
% Plot a simple 2-circle Venn diagram with custom
patch properties
```

```
figure, axis equal, axis off
```

```
A = [40 40]; I = [5];
```

```
venn(A,I,'FaceColor',{'w','w'},'FaceAlpha',{0.6,0.6},
'EdgeColor','black')
```

```
% These diagrams can't be edited properly within
Matlab, but they do
```

```
% provide the basis for editing outside of Matlab.
```

```
gtext('A',FontSize,20);gtext('B',FontSize,20)
```

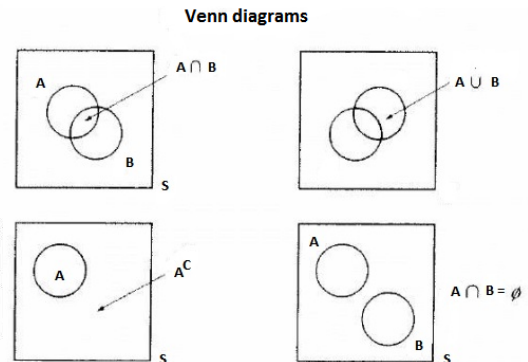


FIGURE 2.2.4

Example 2.2.13 When Swampwater Tech's Class of '55 held its fortieth Reunion, 100 graduates attended ...

```
% LMex020213_4th.m
```

```
% Example 2.2.12 p 33-34 in
```

```
% Larsen & Marx (2006) Introduction to
Mathematical Statistics, 4th edition
```

```
% Written by Eugene.Gallagher@umb.edu 1/14/11
37
```

```
%
```

```
http://alpha.es.umb.edu/faculty/edg/files/edgwebp.ht
ml
```

```
% Plot a simple 3-circle venn diagram with custom
patch properties
```

```
figure, axis equal, axis off
```

```
A = [100 15 30]; I = [15 30 10 10];
```

```
venn(A,I,'FaceColor',{'y','r','g'},'FaceAlpha',{0.4,0.6,
0.6},'EdgeColor','black')
```

```
% These diagrams can't be edited properly within Matlab, but they do
```

```
% provide the basis for editing outside of Matlab.
```

```
% gtext will place the labels A and B
```

```
gtext('S',FontSize,20);gtext('L',FontSize,20),
```

```
gtext('H',FontSize,20)
```

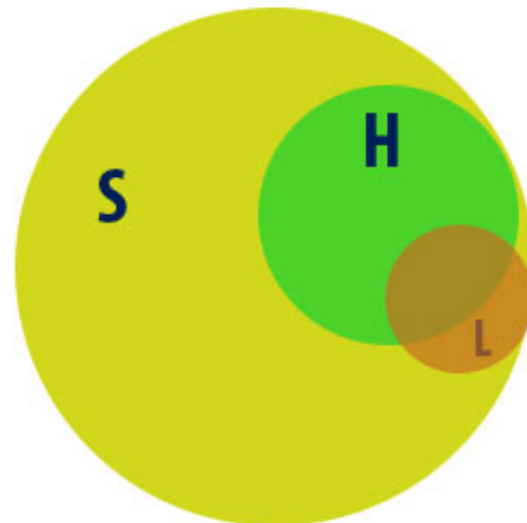


Figure 4.

2.3.31 During orientation week, the latest Batman movie was shown twice at State University. Among the entering class of 6000 freshmen, 850 went to see it the first time, 690 the second time, while 4700 failed to see it either time. How many saw it twice?

2.3.32 Let A and B be any two events. Use Venn diagrams to show that a) the complement of their intersection is the union of their complements:

$$(A \cap B)^c = A^c \cup B^c.$$

the complement of their union is the intersection of their complements:

$$(A \cup B)^c = A^c \cap B^c.$$

These two results are known as DeMorgan's laws.

Let A, B and C be any three events. Use Venn diagrams to show that $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ b) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

Venn diagram

2.3 The Probability Function

2.3.1 Four ways of defining **probability**

2.3.1.1 **Classical probability**

“Imagine an experiment, or game, having n possible outcomes — and suppose that those outcomes are equally likely. If some event A were satisfied by m out of those n , the probability of A [Written $P(A)$] should be set equal to m/n . This is the classical or *a priori* definition of probability.

2.3.1.2 **Empirical probability** (Attributed to von Mises, but can be found at least a century earlier)

“Consider a sample space S , and any event A , defined on A . If our experiment were performed one time, either A or A^c would be the outcome. If it were performed n times, the resulting set of sample outcomes would be members of A on m occasions, m being some integer between 0 and n , inclusive. Hypothetically, we could continue the process an infinite number of times. As n gets large, the ratio m/n will fluctuate less and less. The number that m/n converges to is called the empirical probability of A , that is $P(A) = \lim_{n \rightarrow \infty} m/n$.

2.3.1.3 **Axiomatic probability**. Andrei Kolmogorov

If S has a finite number of members, Kolmogorov showed that as few as three axioms are necessary and sufficient for characterizing the probability function P .

2.3.1.3.1 **Axiom 1.** Let A be any event defined over S . Then $P(A) \geq 0$.

2.3.1.3.2 **Axiom 2** $P(S) = 1$.

2.3.1.3.3 **Axiom 3** Let A and B be any two mutually exclusive events defined over S. Then

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

When S has an infinite number of members, a fourth axiom is needed

2.3.1.3.4 **Axiom 4** Let A_1, A_2, \dots , be events defined over S. If $A_i \cap A_j = \emptyset$ for each $i \neq j$, then

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

From these simple statements, all other properties of the probability function can be derived.

2.3.1.4 Subjective probability

2.3.1.4.1 What is a person's measure of belief that an event will occur?

Some basic properties of P

Theorem 2.3.1 $P(A^c) = 1 - P(A)$

Proof. By axiom 2 and Definition 2.2.3

Theorem 2.3.2 $P(\emptyset) = 0$.

Theorem 2.3.3 If $A \subset B$, then $P(A) \leq P(B)$.

Theorem 2.3.4 For any event A, $P(A) \leq 1$.

Theorem 2.3.5 Let A_1, A_2, \dots, A_n be events defined over S. If $A_i \cap A_j = \emptyset$ for $i \neq j$, then

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

Theorem 2.3.6 $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Example 2.3.1

```
% LMex020301_4th.m
% Example 2.3.1 p 38-39 in
% Larsen & Marx (2006) Introduction to
Mathematical Statistics, 4th edition
% Written by Eugene.Gallagher@umb.edu 1/14/11
%
http://alpha.es.umb.edu/faculty/edg/files/edgwebp.ht
ml
% Plot a simple 2-circle Venn diagram with custom
patch properties
```

```
figure, axis equal, axis off
```

```
A = [.3 .5]; I = [.1];
```

```
% Prints 2.3.1
```

```
% venn(A,I,'FaceColor',{'w','w'},'FaceAlpha',{0.6,0.6},'EdgeColor','black')
```

```
% Prints 2.3.2 with appropriate editing in another program
```

```
venn(A,I,'FaceColor',{'w','g'},'FaceAlpha',{.2,.6},'EdgeColor','black')
```

```
% These diagrams can't be edited properly within Matlab, but they do
```

```
% provide the basis for editing outside of Matlab.
```

```
% gtext will place the labels A and B
```

```
gtext('A',FontSize',20);gtext('B',FontSize',20)
```

```
figure(gcf);pause
```

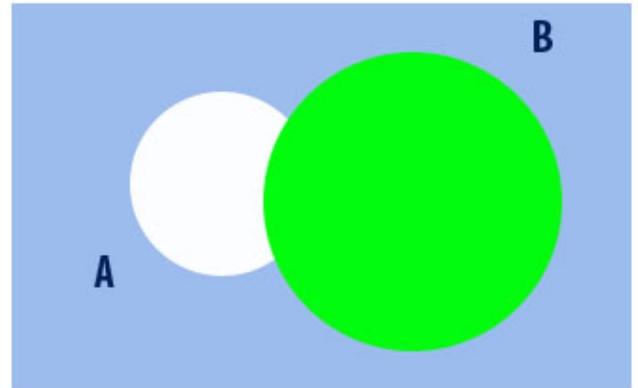


Figure 5. Figure 2.3.2

Example 2.3.2

Show that $P(A \cap B) \geq P(A^c) - P(B^c)$ for any two events A and B defined on a sample space S. Not a Matlab problem

Example 2.3.3

Two cards are drawn from a poker deck without replacement. What is the probability that the second is higher in rank than the first?

Let A_1, A_2, A_3 be the events “First card is lower in rank, “First card is higher in rank, and “Both cards have the same rank,” respectively. Clearly, the three A_i 's are mutually exclusive and they account for all possible outcomes, so from Theorem 2.3.5,

$$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) = P(S) = 1.$$

Once the first card is drawn, there are only three choices for the second that would have the same rank – that is, $P(A_3) = 3/51$. Symmetry, of course, demands that $P(A_1) = P(A_2)$. Therefore,

$$2P(A_2) + \frac{3}{51} = 1.$$

implying that $P(A_2) = \frac{8}{17}$.

```
% LMex020303_4th.m
% Example 2.3.3 from Larsen & Marx (2006,
p 39).
% Written by E. Gallagher 8/4/01; revised
1/14/11
% Written by Eugene.Gallagher@umb.edu
1/14/11
%
http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html
% You draw two cards at random. What is the probability that the
% second card is greater than the first?
% Solved via Monte Carlo simulation.
fprintf('The expected value is %6.4f\n\n',8/17)
% Monte Carlo simulation
deck=1:13;
deck= repmat(deck',1,4); % tile 4 times for the 4 suits
deck=deck(:);
trials=1e5;
twocards=zeros(trials,2); % initialize a matrix to store results
for i=1:trials
    j=randperm(52); % This shuffles the numbers 1 to 52
    % saves the 1st 2 cards after randperm has shuffled the deck
    twocards(i,:)=[deck(j(1)) deck(j(2))];
end
i=find(twocards(:,2)>twocards(:,1)); % compares columns in each row
p=length(i)/trials;
% Confidence intervals will be introduced in Chapter 6:
fprintf('In %6.0f trials, P(2nd card > first) was %6.4f\n',trials, p)
% 95% percent confidence interval for p:
halfwidth=1.96*sqrt(p*(1-p)/trials);
fprintf(...
    'The upper and lower confidence limits are %6.4f
and %6.4f\n',...
    p-halfwidth,p+halfwidth)
```

The expected value is 0.4706

In 100000 trials, P(2nd card > first) was
0.4694

The upper and lower confidence limits are
0.4663 and 0.4725

Example 2.3.4

Martial arts film. The actress appears in 40% of the film's scenes. Her double appears in 30%, and the two of them appear together 5% of the time. What is the probability that in a given scene (a) only the stunt double appears and (b) neither the lead actress nor the double appears?

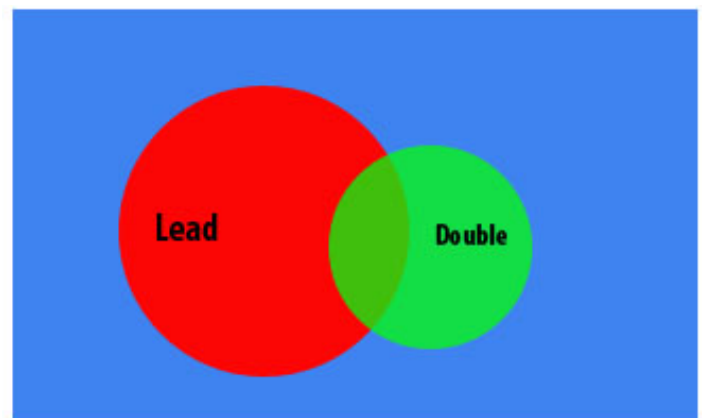


Figure 6. Example 2.3.4

$$\begin{aligned}
 P(\text{Only double appears}) &= P(D) - P(L \cap D) \\
 &= 0.30 - 0.05 \\
 &= 0.25
 \end{aligned}$$

$$\begin{aligned}
 P(\text{Neither appears}) &= 1 - P(L \cup D) \\
 &= 1 - [P(L) + P(D) - P(L \cap D)] \\
 &= 1 - [0.40 + 0.30 - 0.05] \\
 &= 0.35
 \end{aligned}$$

Example 2.3.5 Biff decides to test the medical school waters and sends his MCATs to two colleges, X and Y. Based on how his friends fared, he estimates that his probability of being accepted at X is 0.7 and at Y is 0.4. He also expects that there is a 75% chance that at least one of his applications will be rejected. What is the probability that he gets at least one acceptance. Let A be the event “School A accepts him” and B, the event “School Y accepts him.” We are given that $P(A)=0.7$, $P(B)=0.4$, and $P(A^c \cup B^c)=0.75$. What we are looking for is $P(A \cup B)$.

From Theorem 2.3.6,

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

Recall from Question 2.2.17 that $A^c \cup B^c = (A \cap B)^c$, so

$$P(A \cup B) = 1 - [(P(A) \cap P(B))^c] = 1 - 0.75 = 0.25.$$

$$\begin{aligned}
 P(A \cup B) &= 0.7 + 0.4 - 0.25 \\
 &= 0.85
 \end{aligned}$$

Questions p 41

Violence on TV

2.3.8 Venn diagram asked of Summer 2011 students

2.3.9 Fall 2010 Submit a Matlab m.file that solves 2.3.9 (p. 41) [Hint: The solution is in the back of the book (p. 839). Your program should consist of just 2 lines, but add your documentation with % statements]

Let A and B be any two events defined on S. Suppose that $P(A) = 0.4$, $P(B)=0.5$, and $P(A \cap B) = 0.1$. What is the probability that A or B but not both occur?

2.3.10 To be asked as a Summer 2011 Advanced problem An urn contains 24 chips, numbered 1 through 24. One is drawn at random. Let A be the event that the number is divisible by 2 and let B be the event that the number is divisible by 3. Find $P(A \cup B)$.

2.4 CONDITIONAL PROBABILITY p 42

- 2.4.1 Any probability that is revised to take into account the (known) occurrence of other events is said to be a **conditional probability**.
- 2.4.2 The symbol $P(A|B)$ — read “the probability of A given B” — is used to denote a conditional probability. Specifically $P(A|B)$ refers to the probability that A will occur given that B has already occurred.”

Definition 2.4.1. Let A and B be any two events defined on S such that $P(B) > 0$. The conditional probability of A , assuming that B has already occurred, is written $P(A|B)$ and is given by

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

$$P(A \cap B) = P(A|B) P(B).$$

Example 2.4.1

```
% LMex020401_4th.m
% A silly program to solve Example 2.4.1 on p 44 from
% Modeled after ex 2.2.6 in
% Larsen & Marx (2006) Introduction to Mathematical Statistics, 4th edition
% Written by Eugene.Gallagher@umb.edu 1/14/11 31
% http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html
% A card is drawn from a poker deck. What is the probability that the card
% is a club, given that the card is a king.
% Kings will be 13's in this deck.
% C is the event card is a club
% K is the event card is a king
deckn= repmat(1:13,1,4);
deckf= repmat('HCDS',1,13);
Ki=find(deckn==13);Ci=find(deckf=='C');
% P(C|K)=P(C^K)/P(K);
PK=length(Ki)/length(deckf);
PCintK=length(intersect(Ki,Ci))/length(deckf);
PCgivenK=PCintK/PK;
fprintf('The conditional probability of a Club, \n')
fprintf('given that the card is a King is %6.4f\n',PCgivenK);
% find the unconditional probability P(C)
PC=length(Ci)/length(deckf);
fprintf('The unconditional probability of a Club is % 6.4f\n',PC)
```

Example 2.4.2 p. 44 The two-boys problem

```
% LMex020402_4th.m
% Example 2.4.2 The two-boys problem page 44 in
% Larsen & Marx (2006) Introduction to Mathematical Statistics, 4th edition
% Written by Eugene.Gallagher@umb.edu 1/14/11 30 (with advanced problem,
% analyzing the problem as a Monte Carlo simulation)
% http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html
% Consider the set of families with 2 children. What is the probability
% that both children are boys, given that at least one child is a boy?
% Use LMex020201_4th.m as a model
gender=['B';'G']
% create the sample space
births=2;
S=repmat(' ',2^births,2);
i=1;
for birth1=1:2
    for birth2=1:2
        S(i,:)= [gender(birth1) gender(birth2)];
        i=i+1;
    end
end
disp('The sample space, S:')
S
% Event A=both children are boys, Event B=at least one child is a boy.
%  $P(A|B)=P(A \cap B)/P(B)=P(A)/P(B)$ ;
[r,c]=size(S);
PA=sum(sum(S=='B')==2)/r
PB=sum(sum(S=='B')==0)/r
fprintf('The probability that A, "both of two children are boys", \n')
fprintf('given B "that at least one child is a boy" is PA/PB = %6.4f\n',...
    PA/PB)
```

Example 2.4.3 Not a Matlab-type problem

Example 2.4.4 Not a Matlab-type problem

Example 2.4.5 Not a Matlab-type problem

Case Study 2.4.1 Not enough information to program

Case Study 2.4.2 Monty Hall problem.

% Montyhall.m from Matlab file exchange

Questions. p. 51

2.4.12 Assigned Fall 2010

2.4.18 Assigned Summer 2011 Conditional probability problem with dice

2.4.3 **Applying conditional probabilities to Higher-Order Intersections** (p 53)

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_n | A_1 \cap A_2 \cap \dots \cap A_{n-1}) \cdot P(A_{n-1} | A_1 \cap A_2 \cap \dots \cap A_{n-2}) \cdot \dots \cdot P(A_2 | A_1) \cdot P(A_1). \quad (2.4.3)$$

Example 2.4.6 Not a Good Matlab problem

Case Study 2.4.3 P(Habitable planet)

Questions p. 56

2.4.4 Calculating “Unconditional” Probabilities (p 56)

Theorem 2.4.1 Let $\{A_i\}_{i=1}^n$ be a set of events defined over S such that

$$S = \bigcup_{i=1}^n A_i, \quad A_i \cap A_j = \emptyset \text{ for } i \neq j, \text{ and } P(A_i) > 0 \text{ for } i = 1, 2, \dots, n.$$

For any event B ,

$$P(B) = \sum_{i=1}^n P(B|A_i)P(A_i)$$

Example 2.4.7 Urns

Example 2.4.8 A standard Poker deck is shuffled and the card on top is removed. What is the probability that the second card is an ace? “What you don’t know doesn’t matter”

Example 2.4.9 Ashley’s job interview

Example 2.4.10 Congressional Race

Example 2.4.11

```
% LMex020411_4th.m
% Larsen & Marx Example 2.4.11, page 59 in 4th edition
% 3rd edition Example 2.06.11, page 63 Larsen & Marx, 3rd ed.
% Eugene D. Gallagher, written 1991, revised 1/14/11
% Three chips are placed in an urn. One is Red on both sides, a second
% is blue on both sides and a third is red on one side and blue on the
% other. One chip is selected at random and placed on a table. It is
% Red What is the probability that the color underneath is also red?
% Monte Carlo simulation of process:
coins=['RR';'RB';'BB'];
numredt=0;
numredb=0;
trials=1e3;
for i=1:trials
    pick_coin=randi(3); % picks 1, 2, or 3 randomly
    pick_face=randperm(2); % permutes [1 2] randomly
```

```
if coins(pick_coin,pick_face(1))=='R'  
    numredt=numredt+1;  
    if coins(pick_coin,pick_face(2))=='R'  
        numredb=numredb+1;  
    end  
end  
end  
P=numredb/numredt  
Exp=2/3;  
fprintf('Expected P=%6.4f and a Monte Carlo simulation with %7.0f \n',...  
    Exp, trials)  
fprintf('trials found P = %6.4f\n',P)
```

Questions p 60

2.4.32 Assigned Fall 2010

2.4.36 jury

2.4.5 Bayes Theorem p. 63

Theorem 2.4.2 (Bayes)

Let $\{ A_i \}_{i=1}^n$ be a set of n events, each with positive probability,

that partitions S in such a way that $\bigcup_{i=1}^n A_i = S$ and $A_i \cap A_j = \emptyset$ for $i \neq j$.

For any event B (also defined on S), where $P(B) > 0$,

$$P(A_j|B) = \frac{P(B|A_j)P(A_j)}{\sum_{i=1}^n P(B|A_i)P(A_i)}$$

for any $i \leq j \leq n$.

Let $\{ A_i \}_{i=1}^n$ be a set of n events,
 each with positive probability,
 that partitions S in such a way that

$$\bigcup_{i=1}^n A_i = S$$

and $A_i \cap A_j = \emptyset$ for $i \neq j$.

For any event B (also defined on S),
 where $P(B) > 0$,

$$P(A_j|B) = \frac{P(B|A_j)P(A_j)}{\sum_{i=1}^n P(B|A_i)P(A_i)}$$

for any $i \leq j \leq n$.

Example 2.4.12

```
% LMex020412_4th.m
% Example 2.4.14 Bayes theorem p 63-64 in
% Larsen & Marx (2006) Introduction to Mathematical Statistics, 4th edition
% Written by Eugene.Gallagher@umb.edu, revised 1/15/11
% http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html
% Application of Bayes theorem
% Revised 1/15/2011
% Application of Bayes theorem
% A_1 = Coin Came up heads, chip came from urn 1
% A_2 = Coin came up tails, chip came from urn 2
% B = white chip is drawn
PA1 = 2/3;
PA2 = 1-PA1;
PBgivenA1 = 3/7;
PBgivenA2 = 6/9;
% PA2B= Probability of Tails given that the chip was white.
PA2B=PBgivenA2*PA2/(PBgivenA2*PA2+PBgivenA1*PA1);
fprintf('The probability that the coin toss was tails is %6.4f.\n',PA2B)
format rat
disp(PA2B)
format
% Note that the equation could be modified if there were 3 or more
% partitions:
% PA3B=PBgivenA3*PA3/(PBgivenA3*PA3+PBgivenA2*PA2+PBgivenA1*PA1);
```

Example 2.4.13 Polygraph

```
% LMex020413_4th.m
% Example 2.4.13 Bayes theorem & polygraph p 64-65 in
% Larsen & Marx (2006) Introduction to Mathematical Statistics, 4th edition
% Written by Eugene Gallagher
% Eugene.Gallagher@umb.edu
% Revised 1/3/2010
% Application of Bayes theorem
% P (Suspect innocent | Polygraph says guilty) = P
% P (Polygraph says guilty | Suspect innocent) = .02 % False positive
% P (Suspect innocent) = 1-12/100 =PI
% P (Polygraph says guilty | suspect guilty) = .9
% A1 = Suspect innocent
% PA2B= Probability of Tails given that the chip was white.
P=BayesTheorem(.02,1-12/100,.9) % Calls a function
P2=.02*(1-12/100)/(0.02*(1-12/100)+.9*12/100)
```

Example 2.4.14 Diagnostic test for cancer

```
% LMex020414_4th.m
% Example 2.4.14 Bayes theorem & cancer screening p 65-66 in
% Larsen & Marx (2006) Introduction to Mathematical Statistics, 4th edition
% Written by Eugene Gallagher 2001
% Eugene.Gallagher@umb.edu
% http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html
% Revised 1/15/2011
% Application of Bayes theorem
% PA_1 = Probability of cancer in population = Prevalence Rate
% PA_2 = 1-PA_1=Probability that patient is cancer-free
% PB = Test is positive for cancer
% PBgivenA1 = 0.90 = Sensitivity of the test
% PBgivenA2 = 0.001 = False Positive Frequency
% PBcgivenA1 = False Negative Frequency
% PBcgivenA2 = Screened negative and did not have the disease = Specificity
% Bayes Theorem for a test's yield or predictive value
% Yield = Predictive Value = (Prevalence)*(Sensitivity)/
% (Prevalence)*(Sensitivity)+ (1-Prevalence)*(1-Sensitivity)
PA_1=0.0001
PA_2=1-PA_1
PBgivenA1=0.90
% False positive rate:
PBgivenA2=0.001
PA1givenB=PBgivenA1*PA_1/(PBgivenA1*PA_1+PBgivenA2*PA_2)
disp(' Table 2.4.3')
disp(' P(A_1) P(B|A_1^c) P(A_1|B)')
for PA_1=logspace(-4,-2,3)
  for PBgivenA2=logspace(-3,-4,2)
```

```

PA1givenB=PBgivenA1*PA_1/(PBgivenA1*PA_1+PBgivenA2*PA_2);
disp([PA_1 PBgivenA2 PA1givenB]);
end
end

```

P(A ₁)	P(B A ₁ ^c)	P(A ₁ B)
Disease Prevalence	False Positive Value (Yield)	Predictive Value (Yield)
0.0001	0.0010	0.0826
0.0001	0.0001	0.4737
0.0010	0.0010	0.4737
0.0010	0.0001	0.9000
0.0100	0.0010	0.9000
0.0100	0.0001	0.9890

Example 2.4.15 Home burglar alarm

```

% LMex020415_4th.m
% Example 2.4.15 Bayes theorem & burglar alarms p 66-67 in
% Larsen & Marx (2006) Introduction to Mathematical Statistics, 4th edition
% Written by Eugene Gallagher 2001
% Eugene.Gallagher@umb.edu
% http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html
% Revised 1/15/2011
% Application of Bayes theorem
% PA_1 = Probability of burglaries in population = Prevalence Rate
% PA_2 = 1-PA_1=Probability that patient is cancer-free
% PB = Burglar alarm goes off
% PBgivenA1 = 0.95 = Sensitivity of the test
% PBgivenA2 = 5/370 = False Positive Frequency
% PBcgivenA1 = False Negative Frequency
% PBcgivenA2 = No alarm and no burglary = Specificity
% Bayes Theorem for a test's yield or predictive value
% Yield = Predictive Value = (Prevalence)*(Sensitivity)/
% (Prevalence)*(Sensitivity)+ (1-Prevalence)*(1-Sensitivity)
PA_1=2/1e4
PA_2=1-PA_1
PBgivenA1=0.95
% False positive rate:
PBgivenA2=5/370
PA1givenB=PBgivenA1*PA_1/(PBgivenA1*PA_1+PBgivenA2*PA_2)

```

```
PBgivenA1=[.95 .97 .99 .999];
PA1givenB=(PBgivenA1*PA_1)./(PBgivenA1*PA_1+PBgivenA2*PA_2);
disp('      Table 2.4.4')
disp('      P(B|A_1) ')
disp('      Sensitivity')
fprintf('      %4.2f %4.2f %4.2f %6.4g\n',PBgivenA1)
fprintf('Yield %5.3f %5.3f %5.3f %5.3f\n',PA1givenB)
```

Table 2.4.4				
	P(B A_1)			
	Sensitivity			
	0.95	0.97	0.99	0.999
Yield	0.027	0.028	0.028	0.028

Example 2.4.16 Jeremy & Emma

Questions p. 68-69 2.4.42

2.4.46 Murder she wrote. Assigned Summer 2011

2.4.48 Assigned Fall 2010

2.5 Independence, p. 69

Definition 2.5.1 Two events are said to be **independent** if $P(A \cap B) = P(A) \cdot P(B)$.

Example 2.5.1 Let A be the event of drawing a king from a standard poker deck and B, the event of drawing a diamond. The, by Definition 2.7.1, A and B are independent because the probability of their intersection — drawing a king of diamonds— is equal to $P(A) \cdot P(B)$

Example 2.5.2 Suppose that A and B are independent events. Does it follow that A^c and B^c are also independent? That is, does $P(A \cap B) = P(A) \cdot P(B)$ guarantee that $P(A^c \cap B^c) = P(A^c) \cdot P(B^c)$?
Yes.

Example 2.5.3

```
% LMex020503_4th.m
% Example 2.5.3 Application of solve p 63-64 in
% Larsen & Marx (2006) Introduction to Mathematical Statistics, 4th edition
% Written by Eugene.Gallagher@umb.edu, revised 1/15/11
% http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html
%      Table 2.5.1
%      White Black
% Male   50   30
% Female 40   x
% P(Black ^ Female) = P(Black) * P(Female)
% x/(120+x) = (x+30)/(120+x) * (40+x)/(120+x)
syms x
solve(x/(120+x)-(40+x)/(120+x)*(30+x)/(120+x))
```

Example 2.5.4

Mutually exclusive events aren't necessarily independent.

2.5.1 Deducing independence

Example 2.5.5 Proofreading errors. Simple calculation

Example 2.5.6. Probability of a recessive allele = $(r+q/2)^2$

Example 2.5.7 Probability of different blood types. An example of the Gini-Simpson index of gene heterogeneity.

```
% LMex020507_4th.m
```

```
% Example 2.5.5 Application of Gini-Simpson diversity p 74 in
```

```
% Larsen & Marx (2006) Introduction to Mathematical Statistics, 4th edition
```

```
% Written by Eugene.Gallagher@umb.edu, revised 1/15/11
```

```
DATA=[.4 .1 .05 .45];
```

```
[r,c]=size(DATA);
```

```
N=sum(DATA)';
```

```
P=1-sum(((DATA./repmat(N,1,c)).^2)'); % This formula works for multiple
```

```
% rows of data
```

```
G=gini(DATA); % application of the Gini-Simpson index.
```

```
fprintf(...
```

```
    'Probability that Emma & Josh have different blood types = %5.3f\n',P)
```

```
fprintf('Gini-Simpson index of gene heterogeneity = %5.3f\n',G)
```

```
function G=gini(DATA)
```

```
% Calculate Gini-Simpson's diversity
```

```
% format G=gini(DATA)
```

```
% Also called gene heterogeneity or gene heterozygosity
```

```
% input: a sample x species or sample by character matrix.
```

```
% output: G, a column vector with Gini-Simpson diversities
```

```
% formulas  $G=1-\sum((n_i/N)^2)$ 
```

```
% where N is the sample total and  $n_i$  is the abundance of
```

```
% the  $i$ 'th species.
```

```
% Hurlbert SM (1971) The non-concept of species diversity: a
```

```
% critique and alternative parameters. Ecology 52: 577-586.
```

```
% Magurran AE (1988) Ecological diversity and its measurement.
```

```
% Princeton University Press.
```

```
% written 4/5/94 by E. Gallagher, Environmental Sciences Program
```

```
% UMASS/Boston, Eugene.Gallagher@umb.edu
```

```
% last revised: 1/15/2011
```

```
% See Simpson.m for the unbiased formula
```

```
[r,c]=size(DATA);
```

```
N=sum(DATA)';
```

```
G=1-sum(((DATA./repmat(N,1,c)).^2)');
```

Questions p 74-75

2.5.2 Defining the independence of more than 2 events

Definition 2.7.2 Events A_1, A_2, \dots, A_n are said to be independent if for every set of indices i_1, i_2, \dots, i_k between 1 and n , inclusive,

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \dots \cdot P(A_{i_k})$$

Example 2.5.8 Getting bumped from classes (Some similarities to Hurlbert's $E(S_n)$)

Example 2.5.9 Probability of exactly one death. A Monte Carlo simulation

```
% LMex020509_4th.m
% Example 2.5.9 Application of a Monte Carlo simulation p 63-64 in
% Larsen & Marx (2006) Introduction to Mathematical Statistics, 4th edition
% Written by Eugene.Gallagher@umb.edu, revised 1/15/11
% http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html
% Probability of survival of 3 independent individuals
PA1=.7;PA2=.9;PA3=0.3;
% The probability of exactly one death is
P=PA1*PA2*(1-PA3)+PA1*(1-PA2)*PA3+(1-PA1)*PA2*PA3;
deaths=0;
trials=1e6;
exactlyone=0;
for i=1:trials
    death=0;
    if rand>PA1;death=death+1;end
    if rand>PA2;death=death+1;end
    if rand>PA3;death=death+1;end
    if death==1;
        exactlyone=exactlyone+1;
    end
end
Pexactlyone=exactlyone/trials;
fprintf('The exact P is %5.3f and Monte Carlo simulated P is %5.3f.\n',...
    P,Pexactlyone)
```

Example 2.5.10 Budget approval circuit Protocol for making financial decisions in certain corporations follows the “circuit” pictured in Figure 2.5.1

2.5.3 Repeated Independent Events

Example 2.5.11 Christmas tree lights

Example 2.5.12 Coin tossing example with Bayes theorem

A box contains one two-headed coin and eight fair coins. One is drawn at random and tossed seven times. Suppose that all seven tosses come up heads. What is the probability that the coin is fair?

This is basically a Bayes problem, but the conditional probabilities on the right hand side of **Theorem 2.6.2** appeal to the notion of independence as well. Define the events B:

seven heads occurred in seven tosses; A_1 : coin tossed has two heads; A_2 : coin tossed was fair.

The question is asking for $P(A_2|B)$.

$$P(A_1) = 1/9 \text{ and } P(A_2) = 8/9$$

$$P(B|A_1) = P(\text{Head on 1st toss} \cap \dots \cap \text{Head on 7th toss} | \text{Coin has two heads}) \\ = 1^7 = 1$$

$$P(B|A_2) = (1/2)^7$$

Substitute into **Bayes theorem**

$$\begin{aligned} P(A_2|B) &= \frac{P(B|A_2)P(A_2)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2)} \\ &= \frac{\left(\frac{1}{2}\right)^7 \left(\frac{8}{9}\right)}{1 \left(\frac{1}{9}\right) + \left(\frac{1}{2}\right)^7 \left(\frac{8}{9}\right)} \\ &= 0.06 \end{aligned}$$

Example 2.5.13 Pete Rose hitting safely in 44 consecutive games: $P = (1 - .7^4)^{44}$

Example 2.5.14 Infant mortality

Example 2.5.15 Craps

```
% crapstableMOF65, a script m.file
% Calculates a table of expected values for winning at craps.
% Written by E. Gallagher 6/2/03 for ECOS601, now EEOS601, revised 1/15/11
% Eugene.Gallagher@umb.edu
% http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html
% calls crapsmof65.m, Mosteller's elegant solution to craps odds
Pexact=zeros(1,11);
Pnumber=zeros(1,11);
for i=1:11;
    [Pexact(i),Pnumber(i)]=crapsmof65(i+1); % calculates exact p and
        % probability of each number on
        % 1st toss
end
disp('Column 1: Number rolled on 1st toss');
disp('Column 2: Probability of rolling the number in the 1st column');
disp('Column 3: P(winning) if you roll the number in the 1st column');
disp('Column 4: Column 2 * Column 3')
disp([2:12; Pnumber; Pexact;Pnumber.*Pexact]')
disp('Probability of winning at craps is sum 2nd column times the third')
disp('(= sum of 4th column). Pwinning=sum(Pnumber.*Pexact)')
Pwinning=sum(Pnumber.*Pexact)
function [Pexact,Pnumber,Odds]=crapsmof65(number,H)
% format [Pexact,Pnumber,Odds]=crapsmof65(number,H)
```

```
% input: number=value of 1st roll
% H ='H' for Pnumber & odds of making 4, 6, 8, or 10 the hard way
% output: Pexact= Exact probability of winning, given 1st roll = number
% Pnumber=Probability of rolling that number by rolling 2 dice.
% Odds =Odds of making number = Pexact/(1-Pexact), NaN if
% Pexact=1
% Using the 'method of reduced sample space' from Mosteller, F. 1965. Fifty
% challenging problems in probability with solutions, p. 9
% see craps.m for a more laborious solution using absorbing Markov chains
% written 6/2/03 by E. Gallagher for EEOS601, revised 1/12/11;
% http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html
if nargin > 1
    if H=='H'
        if mod(number,2) | number ==2 | number==12
            error(...
                'Can only make even numbers between 4 & 10 the "hard" way')
            return
        end
    end
end
die=1:6; % create a sample space of two dice rolls
die1= repmat(die,6,1); % create a matrix, with each row equal to 1 to 6;
die= die'; % transpose of die, now a column vector
die2= repmat(die,1,6); % create a matrix, with each column equal to 1 to 6;
sumdice=die1+die2; % This gives the sample space, the 36 sums of two die.
i=find(sumdice==number); % find the indices for each value equal to number;
lengthi=length(i); % find how many sums of 2 die equal the number in input;
if lengthi==0 % only true if the number isn't in the sample space.
    disp('Your number isn't possible, not in the sample space')
    number
    % The following odd syntax will print the number
    fprintf('Your number, %6.0f, isn't possible, not in the sample space\n',number)
    return % This stops the program after displaying the message;
end
[r,c]=size(sumdice); % find size of sample space
Pnumber=lengthi/(r*c); % What is the probability of rolling your number with 1 roll
% of two dice
% Is the craps shooter an immediate winner?
if (number==7 | number==11) % 2 logical statements using the | OR command.
    Pexact=1; % Matlab performs statements in this if section only
    % if the logical statement number equals 7 or equals 11
    % is true.
    Odds=NaN;
    return
end
```



```
% Is the craps shooter an immediate loser?
if (number==2 | number==3 | number==12)
    Pexact=0;
    Odds=0;
    return % exit the function and return to the command window.
end
% Exact probability of making a point - no need for sum of geometric series
% or Markov chains
% Use Mosteller's 'Method of reduced sample space:
% all we need is Probability of rolling a given
% number - which wins - and the probability of rolling a 7, which loses.
% We don't need any of the other probabilities
% (e.g., probability of rolling 2, 3 ...)
% First find probability of rolling a 7, which results in a loss.
j=find(sumdice==7);P7=length(j)/(r*c); % P7 is the probability of rolling 7
if nargin>1 & H=='H'
    ih=find(sumdice==number & die1==number/2);
    lengthih=length(ih);
    Pnumberh=lengthih/(r*c); % Another way to lose is to roll a soft number
    Pexact=Pnumberh/(P7+Pnumber); % You win or lose if you roll an 8
else
    Pexact=Pnumber/(Pnumber+P7);
end
Odds=Pexact/(1-Pexact);
```

Table 2.8.2 Craps	
Winning event, A_i	$P(A_i)$
A_4	$\frac{1}{36}$
A_5	$\frac{16}{360}$
A_6	$\frac{25}{396}$
A_8	$\frac{25}{396}$
A_9	$\frac{16}{360}$
A_{10}	$\frac{1}{36}$
	0.493

Questions p. 83-85

2.5.12 Assigned Fall 2010

2.5.28. The unsinkable tub is sinking.

A Coast Guard dispatcher receives an SOS from a ship that has run aground off the shore of a small island. Before the captain can relay her exact position, though, her radio goes dead. The dispatcher has n helicopter crews he can send out to conduct a search. He suspects the ship is somewhere either south in area I (with probability p) or north in area II (with probability $1-p$). Each of the n rescue parties is equally competent and has probability r of locating the ship given that it has run aground in the sector being searched. How should the dispatcher deploy the helicopter crews to maximize the probability that one of them will find the missing ship? Hint: Assume that m search crews have been sent to area I and $n-m$ are sent to area II. Let B denote the event that the ship is found, let A_1 be the event that the ship is in area I, and let A_2 be the event that the ship is in area II. Use Theorem 2.6.1 to get an expression for $P(B)$; then differentiate with respect to m .

[See **Nahin problem 16**: the unsinkable sub]

2.6 **COMBINATORICS**

2.6.1 Gottfried Wilhelm Leibniz (1646-1716) 1666 treatise, "*Dissertatio de arte cominatoria*" was perhaps the first monograph written on the subject."

2.6.2 **Counting ordered sequences: the multiplication rule**

Multiplication rule If operation A can be performed in m different ways and operation B in n different ways, the sequence (*Operation A, Operation B*) can be performed in $m \cdot n$ ways.

Example 2.6.1 Lock combinations

Example 2.6.2 Bertillon configuration

Example 2.6.3 Braille

Example 2.6.4 March Madness

Example 2.6.5 Piano keys

2.6.3 **Problem-solving hints**

2.6.3.1 Draw a diagram

2.6.3.2 Do a simple-case enumeration

2.6.3.3 If the outcomes to be counted fall into structurally different categories, the total number of outcomes will be the sum (not the product) of the number of outcomes in each category.

Questions p. 91-93

2.6.1 chemical engineer 3 factors

ciphers

$\text{expand}('(a+b+c)*(d+e+f)*(x+y+u+v+w)')$

2.6.4

2.6.5

2.6.6

2.6.7

2.6.8

- 2.6.9
- 2.6.10
- 2.6.11
- 2.6.12
- 2.6.13
- 2.6.14

Counting permutations (when the objects are all distinct)

Theorem 2.9.1 The number of **permutations** of length k that can be formed from a set of n distinct elements, repetitions not allowed, is denoted by the symbol ${}_n P_k$, where

$${}_n P_k = n(n-1)(n-2)\dots(n-k+1) = \frac{n!}{(n-k)!}.$$

Corollary The number of ways to permute an entire set of n distinct objects is n!

Example 2.6.6 [Note genetic implications] How many permutations of length k=3 can be formed from the set of n = 4 distinct elements A, B, C, and D? [24]

```
%LMex020606_4th.m
% Example 2.6.6 p 94 in
% Larsen & Marx (2006) Example 2.2.2, page 24 Figure 2.2.1
% Written by Eugene Gallagher, Eugene.Gallagher@umb.edu
% http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html
% Written 1/15/2011 & last revised 1/15/2011
% How many permutations of length k=3 can be formed from the set of n=4
% distinct elements, A, B, C, and D?
code=perms('ABCD')
perm4=perms('ABCD')
size(unique(perm4,'rows'))
% Table 2.6.2 is generated by deleting the last column.
perm3=perm4(:,1:3)
% Interesting problem, why does deleting the last base make no difference?
size(unique(perm3,'rows'))
```

Example 2.6.7 Browning permutations

Example 2.6.8 Puzzles

Example 2.6.9 p 95 A deck of 52 cards is shuffled and dealt face up in a row. For how many arrangements will the four aces be adjacent?

```
exp(log(49)+gammaln(5)+gammaln(49))
% or simpler
exp(gammaln(5)+gammaln(50))
1.4599e+064
```

Example 2.6.10 Rook's moves In chess a rook can move vertically and horizontally ... In how many ways can 8 distinct rooks be placed on a chess board so that no two can capture each other?

```
[x=cumprod(1:8).^2;x(8)]
```

Example 2.6.11 Friday the 13

```
% LMex020611_4th.m
% Example 2.6.11 p 98 in
% Larsen & Marx (2006) Introduction to Mathematical Statistics, 4th edition
% Written by Eugene Gallagher, Eugene.Gallagher@umb.edu
% http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html
% Written 1/15/2011 & last revised 1/15/2011
a=exp(gammain(5)+gammain(5))
b=factorial(7)
b=perms('BWZCYAD');[r,c]=size(b)
% Remove the semicolon to list all 5040 orders:
b=[b repmat('X',r,1)];
```

Example 2.6.12 Consider the number of 9 digit numbers that can be formed by rearranging without repetition the integers 1 through 9. For how many of the permutations with the 1 and 2 precede the 3 and 4?

At first glance, this seems to be beyond the scope of **Theorem 2.9.1**. With the help of a symmetry argument, the solution is amazingly simple.

Questions p. 99-100

2.6.15 **Counting permutations (when all objects are not all distinct)**

Theorem 2.6.2 The number of ways to arrange n objects, n_1 being of one kind, n_2 of a second kind, ..., and n_r of an r th kind, is

$$\frac{n!}{n_1! n_2! \dots n_r!}$$

where $\sum_{i=1}^r n_i = n$.

Comment Ratios like $n! / (n_1! n_2! \dots n_r!)$ are called **multinomial coefficients** because the general term in the expansion of

$$(x_1 + x_2 + \dots + x_r)^n$$

is

$$\frac{n!}{n_1! n_2! \dots n_r!} x_1^{n_1} x_2^{n_2} \dots x_r^{n_r}$$

Brillouin's (1951) diversity An information-based diversity statistic, appropriate whenever information-based diversity is being measured from a sample or full census. It is always slightly lower than the Shannon-Wiener H' .

$$\begin{aligned}
 \text{Brillouin's } H &= \frac{1}{N} * \ln \left(\frac{M!}{N_1! N_2! \dots N_s!} \right) \\
 &= \frac{1}{N} * \ln \left(\frac{M!}{\prod_{k=1}^s N_k!} \right)
 \end{aligned}
 \tag{15}$$

where, N = Sample total.

N_k = Abundance of species k .

$\frac{M!}{N_1!N_2! \dots N_s!}$ = No. distinguishable permutations with the N_i 's.

Example 2.6.13 How many ways to arrange 1 nickel, 3 dimes, and 2 quarters?

```
DATA=[1 3 2];
```

```
Permutations=exp(sum(DATA)*brillouin(DATA))
```

```
function [HB,E]=brillouin(DATA)
```

```
% Calculation of Brillouin species diversity.
```

```
% HB=(ln N!-sum(ln ni!))/N
```

```
% format [HB,E]=Brillouin(DATA)
```

```
% Input: DATA is a sample by species matrix,
```

```
% Output: HB is Brillouin sample diversity
```

```
% E is the Evenness index, Pielou's V
```

```
% Hurlbert SM (1971) The non-concept of species diversity: a
```

```
% critique and alternative parameters.
```

```
% Ecology 52: 577-586.
```

```
% Magurran AE (1988) Ecological diversity and its measurement.
```

```
% Princeton University Press.
```

```
% written 4/5/94 by E. Gallagher,
```

```
% Dept of Environmental, Earth & Ocean Sciences
```

```
% UMASS/Boston, Eugene.Gallagher@umb.edu
```

```
% last revised 9/27/2010
```

```
[nsamp,nspec]=size(DATA);
```

```
N=sum(DATA)';
```

```
O=ones(nsamp,nspec);
```

```
% find the number of species.
```

```
S=sum(~(~DATA))';
```

```
lnNfact=gammaln(N+1);
```

```
sumlnnifact=sum(gammaln(DATA+O))';
```

```
HB=(lnNfact-sumlnnifact)/N;
```

```
if nargout>1
```

```
NS=floor(N./S);
```

```
r=N-S.*NS;
```

```
HBMAX=(ones(nsamp,1)./N).* ...
```

```
(lnNfact-((S-r).*gammaln(NS+1))+...
```

```
r.*gammaln(NS+2));  
E=HB./HBMAX;  
end
```

Example 2.6.14 Huygens anagram

```
DATA=[7 5 1 5 1 1 7 4 2 9 4 2 1 2 1 5 5];  
Permutations=exp(sum(DATA)*brillouin(DATA))
```

Example 2.6.15 Expansion of $(1+x^5+x^9)^{100}$ Typo in book

```
DATA=[9 9 5 repmat(1,1,97)];sum(DATA)  
Permutations=exp(sum(DATA)*brillouin(DATA))
```

Permutations =

4.2334e+185 not 485,100

Example 2.6.16 Palindromes

```
% LMex020616_4th.m  
% Example 2.6.16 p 103-104 in  
% Larsen & Marx (2006) Introduction to Mathematical Statistics, 4th edition  
% Writtten by Eugene Gallagher, Eugene.Gallagher@umb.edu  
% http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html  
% Written 1/15/2011 & last revised 1/16/2011  
% How many different ways are there to form palindromes from these 12  
% letters: 4 A's, 6 B's, and 2 C's.  
% Program call's Gallagher's brillouin.m, the brillouin's biodiversity  
% index.  
DATA=[4 6 2]/2;  
Permutations=exp(sum(DATA)*brillouin(DATA))  
% Advanced problem: List all 60!  
function V=palindro(N);  
% Palindrome number generator  
% Format: V=palindro(N)  
% Input:  
% N any integer  
% V a vector of ouput results, with the  
% Output: Solution vector, final element is a palindrome  
% e.g., 1754571, or NaN if numeric accuracy exceeded.  
% Algorithm:  
% if N is a palindrome (e.g., 12321) return V=N  
% if N isn't a palindrome then  
% flip N left to right (e.g., 123 => 321) and  
% add to N. Repeat until a palindrome number  
% is achieved or until N>1e15  
% Written by E. Gallagher 10/10/96.  
% See also palindvp.m for symbolic  
iter=100; % 1e15 will be hit long before this iteration limit  
V=zeros(iter,1);  
V(1,1)=N;
```

```

for i=2:2:iter-1
    V(i,1)=round(eval(fliplr(sprintf('% 16.0f',V(i-1,1)))));
    if V(i,1)==V(i-1,1) | V(i,1)>1e15
% Matlab double precision only 16 significant figures
        if(V(i)>1e15);
            V=NaN;
        else
            V(i)=0;
            V=V(~(~V));
        end
        break
    else
        V(i+1,1)=V(i,1)+V(i-1,1);
    end
end
end

```

Example 2.6.17 Deliveryman's walk

```

% LMex020617_4th.m
% Example 2.6.17 p 104 in
% Larsen & Marx (2006) Introduction to
% Mathematical Statistics, 4th edition
% Writtten by Eugene Gallagher,
% Eugene.Gallagher@umb.edu
%

```

<http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html>

```

% Written 1/15/2011 & last revised 1/16/2011
% A deliveryman is currently at Point X and needs to stop at pont O before
% driving through to point Y (see Figure 2.6.14 on p 104 in Larsen & Marx
% (2006). How many different routes can he take wiout ever going out of his
% way?
% Program call's Gallagher's brillouin.m, the brillouin's biodiversity
% index.
DATA=[9 2;5 3];
N=sum(DATA)';
Permutations=exp(sum(N.*brillouin(DATA)));

```

EXAMPLE 2.6.17

A deliveryman is currently at Point X and needs to stop at Point O before driving through to Point Y (see figure 2.6.14). How many different routes can he take without ever going out of his way?

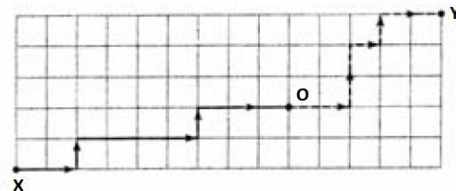


FIGURE 2.6.14

Figure 7. Figure 2.6.14

Questions page 105

2.6.34 Which state can generate more permutations, Tennessee or Florida? Summer 2011 homework problem

2.6.16 Counting combinations

2.6.16.1 We call a collection of unordered elements a **combination** of size k .

Theorem 2.6.3 The number of ways to form **combinations** of size k from a set of n distinct objects, repetitions not allowed, is denoted by the symbols $\binom{n}{k}$ or ${}_n C_k$, where

$$\binom{n}{k} = {}_n C_k = \frac{n!}{(n-k)!}$$

Comment The symbol $\binom{n}{k}$ appears in the statement of the familiar **binomial theorem** from algebra,

$$(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}$$

Since the expression being raised to a power involves two terms, x and y , the constants $\binom{n}{k}$, $k=0, 1, \dots, n$, are commonly referred to as **binomial coefficients** and the expression above is often called Newton's binomial expansion. (**Larsen & Marx 2006 p 111**)

Example 2.6.18

nchoosek(8,2)

Example 2.6.19

nchoosek(6,3)*nchoosek(7,4)*nchoosek(3,2)

Example 2.6.20 p 109

nchoosek(19,2)

Example 2.6.21 p. 109

```
% Example 2.6.21 p 109 in
% Larsen & Marx (2006) Introduction to Mathematical Statistics, 4th edition
% Written by Eugene Gallagher, Eugene.Gallagher@umb.edu
% http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html
% Written 1/15/2011 & last revised 1/16/2011
% Mitch wants to tell 4 jokes at the start of his act and he needs 120
% different combinations of jokes. How many different jokes are needed so
% that he will have at least 120 different opening monologues?
% an application of a while loop and combination function
i=4;joke=0;
display('Table 2.6.3')
display(' n jokes')
while joke<=120
    joke=nchoosek(i,4);
    disp([i joke]);
    i=i+1;
end
```

Example 2.6.22

Proof of the relationship between Pascal's triangle and **binomial coefficients**

Not a Matlab problem

Example 2.6.23

$$\sum_{k=0}^n \binom{n}{k} = 2^n$$

Questions

2.6.54

Combinatorial probability

Example 2.7.1 An urn contains eight chips, numbered 1 through 8. A sample of three is drawn without replacement. What is the probability that the largest chip in the sample is a 5?

% LMex020701_4th.m

% Example 2.6.21 p 109 in

% Larsen & Marx (2006) Introduction to Mathematical Statistics, 4th edition

% Writtten by Eugene Gallagher, Eugene.Gallagher@umb.edu

% <http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html>

% Written 1/15/2011 & last revised 1/16/2011

% An urn contains eight chips, numbered 1 through 8. A sample of three is

% drawn without replacement. What is the probability that the largest chip

% in the sample is a 5?

P=nchoosek(4,2)/nchoosek(8,3)

Example 2.7.2 Not a ready Matlab problem, other than the ratio

Example 2.7.3. Somewhat complicated dice problem; not programmed

Example 2.7.4 Somewhat complicated dice problem; not programmed

Example 2.7.5 TARZAN LIKE JANE

% LMex020705_4th.m

% Example 2.7.5 p 116-117 in

% Larsen & Marx (2006) Introduction to Mathematical Statistics, 4th edition

% Writtten by Eugene Gallagher, Eugene.Gallagher@umb.edu

% <http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html>

% Written 1/16/2011 & last revised 1/16/2011

% Tarzan gives Cheetah the following letters

% AAA EE I J K L NN R T Z

% What is the probability that Cheetah arranges the letters to spell

% TARZAN LIKE JANE, ignore spaces

% using multinomial coefficients

DATA = [3 2 2 1 1 1 1 1 1 1];

format rat

1/exp(sum(DATA)*brillouin(DATA))

% or

exp(gammaln(4)+2*gammaln(3)-gammaln(15))

format

Example 2.7.6 The birthday problem

% LMex020706_4th.m

% Example 2.7.6 The birthday problem p 118-119 in

% Larsen & Marx (2006) Introduction to Mathematical Statistics, 4th edition

```

% Writtten by Eugene Gallagher, Eugene.Gallagher@umb.edu
% http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html
% Written 1/16/2011 & last revised 1/16/2011
% What is the probability that two students out of a group of k will share
% the same birthday?
% calls Gallagher's birthdays.m
disp('      Table 2.7.1')
k=[15 22 23 40 50 70];
fprintf(' k   P(at least 2 with same birthday)\n')
for i=1:length(k)
    P=birthdays(k(i));
    fprintf('%2.0f      %5.3f\n',k(i),P)
end

```

```

function [P,PM]=birthdays(k,trials)
% What is the probability of having two students in a group of k share
% the same birthday?
% format P=birthdays(k,trials)
% input N = number of individuals in group
%   trials (optional) = number of Monte Carlo trials
%   P = exact probability
%   PM = probability based on Monte Carlo simulation
% the birthday problem
% Exact solution & Monte Carlo simulation
% uses Gallagher's ties program for checking birthdays
% Written by E Gallagher, Eugene.Gallagher@umb.edu
% http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html
% Written: 8/5/01, Revised 1/16/11;
P=(365^k - exp(gammainv(366)-gammainv(366-k)))/365^k;
if nargin < 2 & nargin ==2;
    trials=100000; % about 5 seconds
end
if nargin==2
    tally=0;
    for i=1:trials
        draw=randi(365,k,1);
        if length(unique(draw))<k % unique will drop shared birthdays
            tally=tally+1;
        end
    end
    PM=tally/trials;disp(P);
end

```

Example 2.7.7 Poker hands

see http://en.wikipedia.org/wiki/Poker_probability

% LMex020707_4th.m

% Example 2.7.7 Poker hands p 119-121 in

```
% Larsen & Marx (2006) Introduction to Mathematical Statistics, 4th edition
% Written by Eugene Gallagher, Eugene.Gallagher@umb.edu
% http://alpha.es.umb.edu/faculty/edg/files/edgwebp.html
% Written 1/16/2011 & last revised 1/16/2011
% Probability of different hands
% There are nchoosek(52,5) possible hands.
% Full house. nchoosek(13,1) ways to choose denomination and nchoosek(4,3)
% ways of choosing the suit. There are nchoosek(12,1) ways of choosing a
% pair from the denominations remaining (after the denomination of the
% 3-of-a-kind has been selected. There are nchoosek(4,2) ways of choosing
% the suit of the pair.
Pfullhouse=13*nchoosek(4,3)*12*nchoosek(4,2)/nchoosek(52,5)
% one pair
P1pair= 13*nchoosek(4,2)*nchoosek(12,3)*4*4*4/nchoosek(52,5)
% straight; 5 consecutive numbers, but not all in the same suit
Pstraight=(10*4^5-40)/nchoosek(52,5)
% Two pairs, not correct
P2pairs=...
nchoosek(13,2)* nchoosek(4,2) *12*nchoosek(4,2)* nchoosek(11,1)*4/nchoosek(52,5)
P3ofakind=...
13*nchoosek(4,3)*nchoosek(12,2)*4*4/nchoosek(52,5)
```

Problem Solving hints

1. The solution to a combinatorial problem should be set up as a quotient of numerator and denominator *enumerations*.
2. Keep the numerator and denominator consistent with respect to order — if permutations are being counted in the numerator, be sure that permutations are being counted in the denominator; likewise, if the outcomes in the numerator are combinations, the outcomes in the denominator should be combinations
3. The number of outcomes associated with the rolling of n six-sided dice is 6^n ; similarly, the number of outcomes associated with tossing a coin n times is 2^n . The number of outcomes associated with dealing a hand of n cards from a standard 52-card poker deck is ${}_{52}C_n$.

2.7 TAKING A SECOND LOOK AT STATISTICS (ENUMERATION AND MONTE CARLO TECHNIQUES)

2.7.1

- 2.7.2 p. 126 In general efforts to estimate probabilities by simulating repetitions of an experiment (usually with a computer) are referred to as **Monte Carlo studies**.

References

Larsen, R. J. and M. L. Marx. 2006. An introduction to mathematical statistics and its applications, 4th edition. Prentice Hall, Upper Saddle River, NJ. 920 pp. {40}

Index

'at least one'	3, 20, 22
accuracy.	38
Axiomatic probability.	16
Bayes theorem.	3, 5, 6, 24-27, 30, 31
Binomial theorem.	40
Birthday problem.	41, 42
Brillouin's diversity	
definitions.	36
census.	36
Classical probability.	16
combinations.	3, 10, 34, 39, 40, 43
combinatorics.	3, 34
complement.	3, 14, 16
Conditional probability.	3, 6, 21, 22
Confidence interval.	19
Distributions	
empirical.	16
Expected value.	19
Experiment.	3, 7, 8, 11, 16, 43
Fermat.	7
Geometric series.	33
independence.	3, 28-30
intersection.	3, 11-14, 16, 28
Markov chain	
absorbing.	32
Matlab.	3, 5, 7, 9-15, 18, 20, 22, 23, 32, 39-41
Monty Hall problem.	22
multiplication rule.	3, 34
mutually exclusive.	3, 13, 17, 18, 29
Odds.	3, 7, 31-33
OLS (Ordinary Least Squares).	40
Pascal.	7
Pascal's triangle.	40
population.	26, 27
Power.	40
Precision.	39

Probability.	1, 3, 5-7, 9, 10, 16-23, 25-34, 41-43
subjective.	17
Probability function.	16, 17
Quadratic equation.	11
sample.	3, 5, 8-11, 13, 14, 16, 18, 22, 29, 32, 33, 36, 37, 41
Statistic.	1-3, 5, 8, 10-15, 18, 21, 22, 25-30, 36, 38-41, 43, 44
union.	3, 11, 12, 16
Venn diagram.	3, 14-16, 18, 20