Chapter 2
Introduction to Probability

Learning Objectives

1. Obtain an understanding of the role probability information plays in the decision making process.

2. Understand probability as a numerical measure of the likelihood of occurrence.

3. Be able to use the three methods (classical, relative frequency, and subjective) commonly used for assigning probabilities and understand when they should be used.

4. Be able to use the addition law and be able to compute the probabilities of events using conditional probability and the multiplication law.

5. Be able to use new information to revise initial (prior) probability estimates using Bayes' theorem.

6. Know the definition of the following terms:

   experiment
   sample space
   event
   complement
   Venn Diagram
   union of events
   intersection of events
   Bayes' theorem
   addition law
   mutually exclusive
   conditional probability
   independent events
   multiplication law
   prior probability
   posterior probability

Solutions:

1. a. Go to the x-ray department at 9:00 a.m. and record the number of persons waiting.

   b. The experimental outcomes (sample points) are the number of people waiting: 0, 1, 2, 3, and 4.

   Note: While it is theoretically possible for more than 4 people to be waiting, we use what has actually been observed to define the experimental outcomes.

   c.

Number Waiting	Probability
0	.10
1	.25
2	.30
3	.20
4	.15
Total:	1.00

   d. The relative frequency method was used.

2. a. Choose a person at random, have them taste the 4 blends and state their preference.

   b. Assign a probability of 1/4 to each blend. We use the classical method of equally likely outcomes here.

   c.

Blend	Probability
1	.20
2	.30
3	.35
4	.15
Total:	1.00

   The relative frequency method was used.

3. Initially a probability of .20 would be assigned if selection is equally likely. Data does not appear to confirm the belief of equal consumer preference. For example using the relative frequency method we would assign a probability of 5 / 100 = .05 to the design 1 outcome, .15 to design 2, .30 to design 3, .40 to design 4, and .10 to design 5.

4. a. Use the relative frequency approach:

   P(California) = 1,434/2,374 = .60

   b. Number not from 4 states = 2,374 - 1,434 - 390 - 217 - 112 = 221

   P(Not from 4 States) = 221/2,374 = .09

   c. P(Not in Early Stages) = 1 - .22 = .78

   d. Estimate of number of Massachusetts companies in early stage of development - (.22)390 $\approx$ 86

e. If we assume the size of the awards did not differ by states, we can multiply the probability an award went to Colorado by the total venture funds disbursed to get an estimate.

Estimate of Colorado funds = (112/2374)($32.4) = $1.53 billion

Authors' Note: The actual amount going to Colorado was $1.74 billion.

5. a. No, the probabilities do not sum to one. They sum to 0.85.

b. Owner must revise the probabilities so that they sum to 1.00.

6. a. $P(A) = P(150 - 199) + P(200 \text{ and over})$

$$= \frac{26}{100} + \frac{5}{100}$$

$= 0.31$

b. $P(B) = P(\text{less than } 50) + P(50 - 99) + P(100 - 149)$

$= 0.13 + 0.22 + 0.34$

$= 0.69$

7. a. P(A) = .40, P(B) = .40, P(C) = .60

b. $P(A \cup B) = P(E_1, E_2, E_3, E_4) = .80$. Yes $P(A \cup B) = P(A) + P(B)$.

c. $A^c = \{E_3, E_4, E_5\}$ $C^c = \{E_1, E_4\}$ $P(A^c) = .60$ $P(C^c) = .40$

d. $A \cup B^c = \{E_1, E_2, E_5\}$ $P(A \cup B^c) = .60$

e. $P(B \cup C) = P(E_2, E_3, E_4, E_5) = .80$

8. Let Y = high one-year return
M = high five-year return

a. P(Y) = 15/30 = .50

P(M) = 12/30 = .40

$P(Y \cap M) = 6/30 = .20$

b. $P(Y \cup M) = P(Y) + P(M) - P(Y \cap M)$

$= .50 + .40 - .20 = .70$

c. $1 - P(Y \cup M) = 1 - .70 = .30$

9. Let E = event patient treated experienced eye relief.
S = event patient treated had skin rash clear up.

Given:

$P(E) = 90 / 250 = 0.36$

$P(S) = 135 / 250 = 0.54$

$P(E \cup S) = 45 / 250 = 0.18$

$$P(E \cup S) = P(E) + P(S) - P(E \cap S)$$
$$= 0.36 + 0.54 - 0.18$$
$$= 0.72$$

10. P(Defective and Minor) = 4/25

P(Defective and Major) = 2/25

P(Defective) = (4/25) + (2/25) = 6/25

P(Major Defect | Defective) = P(Defective and Major) / P(Defective) = (2/25)/(6/25) = 2/6 = 1/3.

11. a. Yes; the person cannot be in an automobile and a bus at the same time.

b. $P(B^c) = 1 - P(B) = 1 - 0.35 = 0.65$

12. a. $P(A|B) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{0.40}{0.60} = 0.6667$

b. $P(B|A) = \dfrac{P(A \cap B)}{P(A)} = \dfrac{0.40}{0.50} = 0.80$

c. No because $P(A | B) \neq P(A)$

13. a.

	Reason for Applying			
	Quality	Cost/Convenience	Other	Total
Full Time	0.218	0.204	0.039	0.461
Part Time	0.208	0.307	0.024	0.539
Total	0.426	0.511	0.063	1.00

b. It is most likely a student will cite cost or convenience as the first reason: probability = 0.511. School quality is the first reason cited by the second largest number of students: probability = 0.426.

c. P (Quality | full time) = 0.218/0.461 = 0.473
d. P (Quality | part time) = 0.208/0.539 = 0.386

e. $P(B) = 0.426$ and $P(B | A) = 0.473$

Since $P(B) \neq P(B | A)$, the events are dependent.

14.

	\$0-\$499	\$500-\$999	>=\$1000	
<2 yrs	120	240	90	450
>= 2 yrs	75	275	200	550
	195	515	290	1000

	\$0-\$499	\$500-\$999	>=\$1000	
<2 yrs	0.12	0.24	0.09	0.45
>= 2 yrs	0.075	0.275	0.2	0.55
	0.195	0.515	0.29	1.00

a. P(< 2 yrs) = .45

b. P(>= \$1000) = .29

c. P(2 accounts have > = \$1000) = (.29)(.29) = .0841

d. P(\$500-\$999 | >= 2 yrs) = P(\$500-\$999 and >= 2 yrs) / P(>=2yrs) = .275/.55 = .5

e. P(< 2 yrs and >=\$1000) = .09

f. P(>=2 yrs | \$500-\$999) = .275/.515 = .533981

15. a. Total sample size = 2000

Dividing each entry by 2000 provides the following joint probability table.

	Health Insurance		
Age	Yes	No	Total
18 to 34	.375	.085	.46
35 and over	.475	.065	.54
	.850	.150	1.00

Let A = 18 to 34 age group
B = 35 and over age group
Y = Insurance coverage
N = No insurance coverage

b. P(A) = .46
P(B) = .54

Of population age 18 and over

46% are ages 18 to 34
54% are ages 35 and over

c. P(N) = .15

d. $$P(N|A) = \frac{P(N \cap A)}{P(A)} = \frac{.085}{.46} = .1848$$

e. $$P(N|B) = \frac{P(N \cap B)}{P(B)} = \frac{.065}{.54} = .1204$$

f. $P(A|N) = \frac{P(A \cap N)}{P(N)} = \frac{.085}{.150} = .5677$

g. Probability of no health insurance coverage is .15. A higher probability exists for the younger population. Ages 18 to 34: .1848 or approximately 18.5% of the age group. Ages 35 and over: .1204 or approximately 12% of the age group. Of the no insurance group, more are in the 18 to 34 age group: .5677, or approximately 57% are ages 18 to 34.

16. a. $P(A \cap B) = P(A)P(B) = (0.55)(0.35) = 0.19$

b. $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.90 - 0.19 = 0.71$

c. 1 - 0.71 = 0.29

17. a.

	Satisfaction Score					
Occupation	Under 50	50-59	60-69	70-79	80-89	Total
Cabinetmaker	.000	.050	.100	.075	.025	.250
Lawyer	.150	.050	.025	.025	.000	.250
Physical Therapist	.000	.125	.050	.025	.050	.250
Systems Analyst	.050	.025	.100	.075	.000	.250
Total	.200	.250	.275	.200	.075	1.000

b. P(80s) = .075 (a marginal probability)

c. P(80s | PT) = .050/.250 = .20 (a conditional probability)

d. P(L) = .250 (a marginal probability)

e. $P(L \cap \text{Under } 50)$ = .150 (a joint probability)

f. P(Under 50 | L) = .150/.250 = .60 (a conditional probability)

g. P(70 or higher) = .275 (Sum of marginal probabilities)

18. a. $P(B) = 0.25$

$P(S|B) = 0.40$
$P(S \cap B) = 0.25(0.40) = 0.10$

b. $P(B|S) = \frac{P(S \cap B)}{P(S)} = \frac{0.10}{0.40} = 0.25$

c. B and S are independent. The program appears to have no effect.

19. Let: A = lost time accident in current year
B = lost time accident previous year

$\therefore$ Given: P(B) = 0.06, P(A) = 0.05, $P(A|B) = 0.15$

a. $P(A \cap B) = P(A|B)P(B) = 0.15(0.06) = 0.009$

b. $P(A \cup B) = P(A) + P(B) - P(A \mid B)$
$= 0.06 + 0.05 - 0.009 = 0.101$ or 10.1%

20. a. $P(B \cap A_1) = P(A_1)P(B \mid A_1) = (0.20)(0.50) = 0.10$

$P(B \cap A_2) = P(A_2)P(B \mid A_2) = (0.50)(0.40) = 0.20$

$P(B \cap A_3) = P(A_3)P(B \mid A_3) = (0.30)(0.30) = 0.09$

b. $$P(A_2 \mid B) = \frac{0.20}{0.10 + 0.20 + 0.09} = 0.51$$

c.

Events	$P(A_i)$	$P(B \mid A_i)$	$P(A_i \cap B)$	$P(A_i \mid B)$
A_1	0.20	0.50	0.10	0.26
A_2	0.50	0.40	0.20	0.51
A_3	0.30	0.30	0.09	0.23
	1.00		0.39	1.00

21. S_1 = successful, S_2 = not successful and B = request received for additional information.

a. $P(S_1) = 0.50$

b. $P(B \mid S_1) = 0.75$

c. $$P(S_1 \mid B) = \frac{(0.50)(0.75)}{(0.50)(0.75) + (0.50)(0.40)} = \frac{0.375}{0.575} = 0.65$$

22. a. Let F = female. Using past history as a guide, P(F) = .40

b. Let D = Dillard's

$$P(F \mid D) = \frac{.40(3/4)}{.40(3/4) + .60(1/4)} = \frac{.30}{.30 + .15} = .67$$

The revised (posterior) probability that the visitor is female is .67.

We should display the offer that appeals to female visitors.

23. a. P(Oil) = 0.50 + 0.20 = 0.70

b. Let S = Soil test results

Events	$P(A_i)$	$P(S \mid A_i)$	$P(A_i \cap S)$	$P(A_i \mid S)$
High Quality (A_1)	0.50	0.20	0.10	0.31
Medium Quality (A_2)	0.20	0.80	0.16	0.50
No Oil (A_3)	0.30	0.20	0.06	0.19
	1.00		P(S) = 0.32	1.00

P(Oil) = 0.81 which is good; however, probabilities now favor medium quality rather than high quality oil.

24. Let: S = small car
S^c = other type of vehicle
F = accident leads to fatality for vehicle occupant

We have P(S) = .18, so P(S^c) = .82. Also P(F | S) = .128 and P(F | S^c) = .05. Using the tabular form of Bayes Theorem provides:

Events	Prior Probabilities	Conditional Probabilities	Joint Probabilities	Posterior Probabilities
S	.18	.128	.023	.36
S^c	.82	.050	.041	.64
	1.00		.064	1.00

From the posterior probability column, we have P(S | F) = .36. So, if an accident leads to a fatality, the probability a small car was involved is .36.

25.

Events	$P(A_i)$	$P(D \mid A_i)$	$P(A_i \cap D)$	$P(A_i \mid D)$
Supplier A	0.60	0.0025	0.0015	0.23
Supplier B	0.30	0.0100	0.0030	0.46
Supplier C	0.10	0.0200	0.0020	0.31
	1.00		P(D) = 0.0065	1.00

a. P(D) = 0.0065

b. B is the most likely supplier if a defect is found.

26. a.

Events	$P(D_i)$	$P(S_1 \mid D_i)$	$P(D_i \cap S_1)$	$P(D_i \mid S_1)$
D_1	.60	.15	.090	.2195
D_2	.40	.80	.320	.7805
	1.00		$P(S_1)$ = .410	1.0000

$P(D_1 \mid S_1) = .2195$

$P(D_2 \mid S_1) = .7805$

b.

Events	$P(D_i)$	$P(S_2 \mid D_i)$	$P(D_i \cap S_2)$	$P(D_i \mid S_2)$
D_1	.60	.10	.060	.500
D_2	.40	.15	.060	.500
	1.00		$P(S_2)$ = .120	1.000

$P(D_1 \mid S_2) = .50$

$P(D_2 \mid S_2) = .50$

c.

Events	$P(D_i)$	$P(S_3 \mid D_i)$	$P(D_i \cap S_3)$	$P(D_i \mid S_3)$
D_1	.60	.15	.090	.8824
D_2	.40	.03	.012	.1176
	1.00		$P(S_3) = .102$	1.0000

$P(D_1 \mid S_3) = .8824$

$P(D_2 \mid S_3) = .1176$

d. Use the posterior probabilities from part (a) as the prior probabilities here.

Events	$P(D_i)$	$P(S_2 \mid D_i)$	$P(D_i \cap S_2)$	$P(D_i \mid S_2)$
D_1	.2195	.10	.0220	.1582
D_2	.7805	.15	.1171	.8418
	1.0000		.1391	1.0000

$P(D_1 \mid S_1 \text{ and } S_2) = .1582$

$P(D_2 \mid S_1 \text{ and } S_2) = .8418$

Chapter 2
Introduction to Probability

Case Problem: Hamilton County Judges

The data in the table provides the basis for the analysis. We provide notes as a guide to answering questions 1 through 5.

1. The conditional probabilities of cases being appealed in the three courts are given in the 3 Total rows in the table. For Common Pleas Court, the probability of an appeal is .0401; for Domestic Relations Court, the probability of an appeal is .00348; and for Municipal Court, the probability of an appeal is .00461. Appeals are much more likely in Common Pleas Court. But, even there, only 1 in 25 cases are appealed. The unconditional probability of an appeal across all 3 courts is

   (1762 + 106 + 500)/(43,945 + 30,499 + 108,464) = .0129.

2. The probability of a case being appealed for each judge is given in column 5 of the table. Judges Winkler, Panioto and Grady have the lowest probability of appeal for Common Pleas, Domestic Relations and Municipal Courts respectively.

3. The probability of a case being reversed for each judge is given in column 7 of the table. Judges Winkler, Panioto and Grady/Hair have the lowest probability of reversal for Common Pleas, Domestic Relations and Municipal Courts respectively. These are the probabilities for reversal for all cases disposed of, not just the ones appealed.

4. The probability of a reversal given an appeal for each judge is given in column 9 of the table. Judges Nurre, Panioto and Grady/Hair have the lowest probability of reversal for Common Pleas, Domestic Relations and Municipal Courts respectively.

5. We describe here how *The Cincinnati Enquirer* used this data to rank the judges. Other approaches may also be valid, but a rationale should be provided. The newspaper provided ranking for each judge within each of the courts on percentage of cases appealed, percentage of cases reversed and percentage of appealed cases reversed. Those rankings were the same as the ones we have computed based on probabilities in columns 6, 8 and 10 of the table. Then they summed the 3 rankings to come up with a total ranking for each judge. We provide those total ranks in column 11 of the table. Judge Winkler is the highest ranked judge in Common Pleas Court, Judge Panioto is the highest ranked judge in Domestic Relations Court and Judge Grady is the highest ranked judge in Municipal Court.

Common Pleas Court

Judge	Total Cases Disposed	Appealed Cases	Reversed Cases	Probability of Appeal	Rank	Probability of Reversal	Rank	Conditional Probability of Reversal Given Appeal	Rank	Sum of Ranks
Fred Cartolano	3037	137	12	0.04511	14	0.00395	6	0.08759	5	25
Thomas Crush	3372	119	10	0.03529	4	0.00297	4	0.08403	4	12
Patrick Dinkelacker	1258	44	8	0.03498	3	0.00636	12	0.18182	14	29
Timothy Hogan	1954	60	7	0.03071	2	0.00358	5	0.11667	9	16
Robert Kraft	3138	127	7	0.04047	10	0.00223	3	0.05512	2	15
William Mathews	2264	91	18	0.04019	7	0.00795	15	0.19780	16	38
William Morrissey	3032	121	22	0.03991	6	0.00726	14	0.18182	14	34
Norbert Nadel	2959	131	20	0.04427	13	0.00676	13	0.15267	12	38
Arthur Ney, Jr.	3219	125	14	0.03883	5	0.00435	9	0.11200	8	22
Richard Niehaus	3353	137	16	0.04086	11	0.00477	10	0.11679	10	31
Thomas Nurre	3000	121	6	0.04033	8	0.00200	2	0.04959	1	11
John O'Connor	2969	129	12	0.04345	12	0.00404	7	0.09302	6	25
Robert Ruehlman	3205	145	18	0.04524	15	0.00562	11	0.12414	11	37
J. Howard Sundermann Jr.	955	60	10	0.06283	16	0.01047	16	0.16667	13	45
Ann Marie Tracey	3141	127	13	0.04043	9	0.00414	8	0.10236	7	24
Ralph Winkler	3089	88	6	0.02849	1	0.00194	1	0.06818	3	5
Total	43945	1762	199	0.0401		0.00453		0.11294		

Domestic Relations Court

Judge	Total Cases Disposed	Appealed Cases	Reversed Cases	Probability of Appeal	Rank	Probability of Reversal	Rank	Conditional Probability of Reversal Given Appeal	Rank	Sum of Ranks
Penelope Cunningham	2729	7	1	0.00257	2	0.00037	2	0.14286	2	6
Patrick Dinkelacker	6001	19	4	0.00317	3	0.00067	3	0.21053	4	10
Deborah Gaines	8799	48	9	0.00546	4	0.00102	4	0.18750	3	11
Ronald Panioto	12970	32	3	0.00247	1	0.00023	1	0.09375	1	3
Total	30499	106	17	0.00348		0.00056		0.16038		

Municipal Court

Judge	Total Cases Disposed	Appealed Cases	Reversed Cases	Probability of Appeal	Rank	Probability of Reversal	Rank	Conditional Probability of Reversal Given Appeal	Rank	Sum of Ranks
Mike Allen	6149	43	4	0.00699	20	0.00065	7	0.09302	4	31
Nadine Allen	7812	34	6	0.00435	9	0.00077	11	0.17647	10	30
Timothy Black	7954	41	6	0.00515	12	0.00075	10	0.14634	6	28
David Davis	7736	43	5	0.00556	15	0.00065	6	0.11628	5	26
Leslie Isaiah Gaines	5282	35	13	0.00663	19	0.00246	20	0.37143	18	57
Karla Grady	5253	6	0	0.00114	1	0.00000	1	0.00000	1	3
Deidra Hair	2532	5	0	0.00197	3	0.00000	1	0.00000	1	5
Dennis Helmick	7900	29	5	0.00367	6	0.00063	5	0.17241	9	20
Timothy Hogan	2308	13	2	0.00563	17	0.00087	13	0.15385	7	37
James Patrick Kenney	2798	6	1	0.00214	4	0.00036	4	0.16667	8	16
Joseph Luebbers	4698	25	8	0.00532	14	0.00170	18	0.32000	16	48
William Mallory	8277	38	9	0.00459	11	0.00109	14	0.23684	14	39
Melba Marsh	8219	34	7	0.00414	8	0.00085	12	0.20588	13	33
Beth Mattingly	2971	13	1	0.00438	10	0.00034	3	0.07692	3	16
Albert Mestemaker	4975	28	9	0.00563	16	0.00181	19	0.32143	17	52
Mark Painter	2239	7	3	0.00313	5	0.00134	16	0.42857	19	40
Jack Rosen	7790	41	13	0.00526	13	0.00167	17	0.31707	15	45
Mark Schweikert	5403	33	6	0.00611	18	0.00111	15	0.18182	11	44
David Stockdale	5371	22	4	0.00410	7	0.00074	9	0.18182	11	27
John A. West	2797	4	2	0.00143	2	0.00072	8	0.50000	20	30
Total	108464	500	104	0.00461		0.00096		0.20800		