Appendix B–1

Writing the Research Report

After conducting a statistical study, a researcher must write a final report explaining how the study was conducted and giving the results. The formats of research reports, theses, and dissertations vary from school to school; however, they tend to follow the general format explained here.

Front Materials

The front materials typically include the following items:

Title page
Copyright page
Acknowledgments
Table of contents
Table of appendixes
List of tables
List of figures

Chapter 1: Nature and Background of the Study

This chapter should introduce the reader to the nature of the study and present some discussion on the background. It should contain the following information:

Introduction
Statement of the problem
Background of the problem
Rationale for the study
Research questions and/or hypotheses
Assumptions, limitations, and delimitations
Definitions of terms

Chapter 2: Review of Literature

This chapter should explain what has been done in previous research related to the study. It should contain the following information:

Prior research
Related literature

Chapter 3: Methodology

This chapter should explain how the study was conducted. It should contain the following information:

Development of questionnaires, tests, survey instruments, etc.
Definition of the population
Sampling methods used
How the data were collected
Research design used
Statistical tests that will be used to analyze the data

Chapter 4: Analysis of Data

This chapter should explain the results of the statistical analysis of the data. It should state whether the null hypothesis should be rejected. Any statistical tables used to analyze the data should be included here.

Chapter 5: Summary, Conclusions, and Recommendations

This chapter summarizes the results of the study and explains any conclusions that have resulted from the statistical analysis of the data. The researchers should cite and explain any shortcomings of the study. Recommendations obtained from the study should be included here, and further studies should be suggested.

Bayes' Theorem

Objective **B-1**

Find the probability of an event, using Bayes' theorem.

Given two dependent events A and B, the previous formulas for conditional probability allow you to find $P(A$ and $B)$, or $P(B|A)$. Related to these formulas is a rule developed by the English Presbyterian minister Thomas Bayes (1702–1761). The rule is known as **Bayes' theorem.**

It is possible, given the outcome of the second event in a sequence of two events, to determine the probability of various possibilities for the first event. In Example 4–31, there were two boxes, each containing red balls and blue balls. A box was selected and a ball was drawn. The example asked for the probability that the ball selected was red. Now a different question can be asked: If the ball is red, what is the probability it came from box 1? In this case, the outcome is known, a red ball was selected, and you are asked to find the probability that it is a result of a previous event, that it came from box 1. Bayes' theorem can enable you to compute this probability and can be explained by using tree diagrams.

The tree diagram for the solution of Example 4–31 is shown in Figure B–1, along with the appropriate notation and the corresponding probabilities. In this case, A_1 is the event of selecting box 1, A_2 is the event of selecting box 2, R is the event of selecting a red ball, and B is the event of selecting a blue ball.

To answer the question "If the ball selected is red, what is the probability that it came from box 1?" two formulas

$$P(B|A) = \frac{P(A \text{ and } B)}{P(A)} \qquad (1)$$

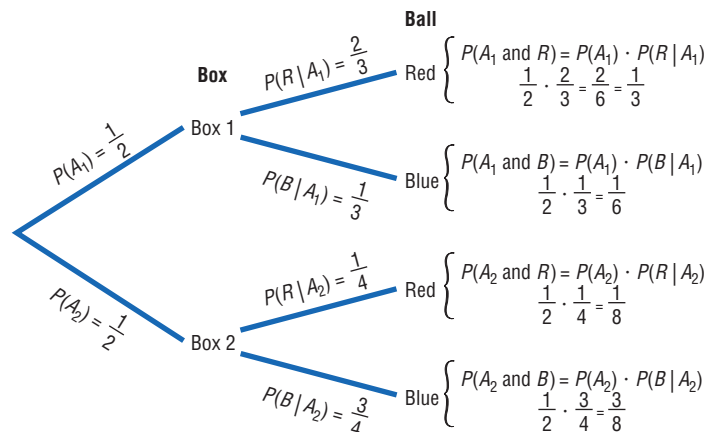$$P(A \text{ and } B) = P(A) \cdot P(B|A) \qquad (2)$$

can be used. The notation that will be used is that of Example 4–31, shown in Figure B–1. Finding the probability that box 1 was selected given that the ball selected was red can be written symbolically as $P(A_1|R)$. By formula 1,

$$P(A_1|R) = \frac{P(R \text{ and } A_1)}{P(R_1)}$$

Note: $P(R \text{ and } A_1) = P(A_1 \text{ and } R)$.

Figure B–1

Tree Diagram for Example 4–31

A–9

By formula 2,

$$P(A_1 \text{ and } R) = P(A_1) \cdot P(R|A_1)$$

and

$$P(R) = P(A_1 \text{ and } R) + P(A_2 \text{ and } R)$$

as shown in Figure B–1; $P(R)$ was found by adding the products of the probabilities of the branches in which a red ball was selected. Now,

$$P(A_1 \text{ and } R) = P(A_1) \cdot P(R|A_1)$$
$$P(A_2 \text{ and } R) = P(A_2) \cdot P(R|A_2)$$

Substituting these values in the original formula for $P(A_1|R)$, you get

$$P(A_1|R) = \frac{P(A_1) \cdot P(R|A_1)}{P(A_1) \cdot P(R|A_1) + P(A_2) \cdot P(R|A_2)}$$

Refer to Figure B–1. The numerator of the fraction is the product of the top branch of the tree diagram, which consists of selecting a red ball and selecting box 1. And the denominator is the sum of the products of the two branches of the tree where the red ball was selected.

Using this formula and the probability values shown in Figure B–1, you can find the probability that box 1 was selected given that the ball was red, as shown.

$$P(A_1|R) = \frac{P(A_1) \cdot P(R|A_1)}{P(A_1) \cdot P(R|A_1) + P(A_2) \cdot P(R|A_2)}$$

$$= \frac{\frac{1}{2} \cdot \frac{2}{3}}{\frac{1}{2} \cdot \frac{2}{3} + \frac{1}{2} \cdot \frac{1}{4}} = \frac{\frac{1}{3}}{\frac{1}{3} + \frac{1}{8}} = \frac{\frac{1}{3}}{\frac{8}{24} + \frac{3}{24}} = \frac{\frac{1}{3}}{\frac{11}{24}}$$

$$= \frac{1}{3} \div \frac{11}{24} = \frac{1}{\overset{1}{\cancel{3}}} \cdot \frac{\overset{8}{\cancel{24}}}{11} = \frac{8}{11}$$

This formula is a simplified version of Bayes' theorem.

Before Bayes' theorem is stated, another example is shown.

Example B–1

A shipment of two boxes, each containing six telephones, is received by a store. Box 1 contains one defective phone, and box 2 contains two defective phones. After the boxes are unpacked, a phone is selected and found to be defective. Find the probability that it came from box 2.

Solution

STEP 1 Select the proper notation. Let A_1 represent box 1 and A_2 represent box 2. Let D represent a defective phone and ND represent a phone that is not defective.

STEP 2 Draw a tree diagram and find the corresponding probabilities for each branch. The probability of selecting box 1 is $\frac{1}{2}$, and the probability of selecting box 2 is $\frac{1}{2}$. Since there is one defective phone in box 1, the probability of selecting it is $\frac{1}{6}$. The probability of selecting a nondefective phone from box 1 is $\frac{5}{6}$.

Since there are two defective phones in box 2, the probability of selecting a defective phone from box 2 is $\frac{2}{6}$, or $\frac{1}{3}$; and the probability of selecting a nondefective phone is $\frac{4}{6}$, or $\frac{2}{3}$. The tree diagram is shown in Figure B–2.

STEP 3 Write the corresponding formula. Since the example is asking for the probability that, given a defective phone, it came from box 2, the corresponding formula is as shown.

$$P(A_2|D) = \frac{P(A_2) \cdot P(D|A_2)}{P(A_1) \cdot P(D|A_1) + P(A_2) \cdot P(D|A_2)}$$

$$= \frac{\frac{1}{2} \cdot \frac{2}{6}}{\frac{1}{2} \cdot \frac{1}{6} + \frac{1}{2} \cdot \frac{2}{6}} = \frac{\frac{1}{6}}{\frac{1}{12} + \frac{2}{12}} = \frac{\frac{1}{6}}{\frac{3}{12}}$$

$$= \frac{1}{6} \div \frac{3}{12} = \frac{1}{\overset{1}{\cancel{6}}} \cdot \frac{\overset{2}{\cancel{12}}}{3} = \frac{2}{3}$$

Figure B–2

Tree Diagram for Example B–1
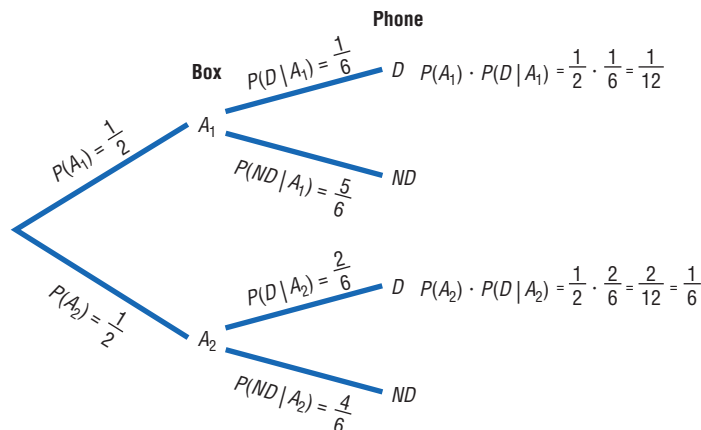
Bayes' theorem can be generalized to events with three or more outcomes and formally stated as in the next box.

Bayes' theorem For two events A and B, where event B follows event A, event A can occur in $A_1, A_2, \ldots, A_n$ mutually exclusive ways, and event B can occur in $B_1, B_2, \ldots, B_m$ mutually exclusive ways,

$$P(A_1|B_1) = \frac{P(A_1) \cdot P(B_1|A_1)}{[P(A_1) \cdot P(B_1|A_1) + P(A_2) \cdot P(B_1|A_2)} \\ {+ \cdots + P(A_n) \cdot P(B_1|A_n)]}$$

for any specific events A_1 and B_1.

The numerator is the product of the probabilities on the branch of the tree that consists of outcomes A_1 and B_1. The denominator is the sum of the products of the probabilities of the branches containing B_1 and A_1, B_1 and A_2, ..., B_1 and A_n.

Example B–2

On a game show, a contestant can select one of four boxes. Box 1 contains one $100 bill and nine $1 bills. Box 2 contains two $100 bills and eight $1 bills. Box 3 contains three $100 bills and seven $1 bills. Box 4 contains five $100 bills and five $1 bills. The contestant selects a box at random and selects a bill from the box at random. If a $100 bill is selected, find the probability that it came from box 4.

Solution

STEP 1 Select the proper notation. Let B_1, B_2, B_3, and B_4 represent the boxes and 100 and 1 represent the values of the bills in the boxes.

STEP 2 Draw a tree diagram and find the corresponding probabilities. The probability of selecting each box is $\frac{1}{4}$, or 0.25. The probabilities of selecting the $100 bill from each box, respectively, are $\frac{1}{10} = 0.1$, $\frac{2}{10} = 0.2$, $\frac{3}{10} = 0.3$, and $\frac{5}{10} = 0.5$. The tree diagram is shown in Figure B–3.

STEP 3 Using Bayes' theorem, write the corresponding formula. Since the example asks for the probability that box 4 was selected, given that $100 was obtained, the corresponding formula is as follows:

$$P(B_4|100) = \frac{P(B_4) \cdot P(100|B_4)}{[P(B_1) \cdot P(100|B_1) + P(B_2) \cdot P(100|B_2)} \\ {+ P(B_3) \cdot P(100|B_3) + P(B_4) \cdot P(100|B_4)]}$$

$$= \frac{0.125}{0.025 + 0.05 + 0.075 + 0.125}$$

$$= \frac{0.125}{0.275} = 0.455$$

Figure B–3

Tree Diagram for Example B–2
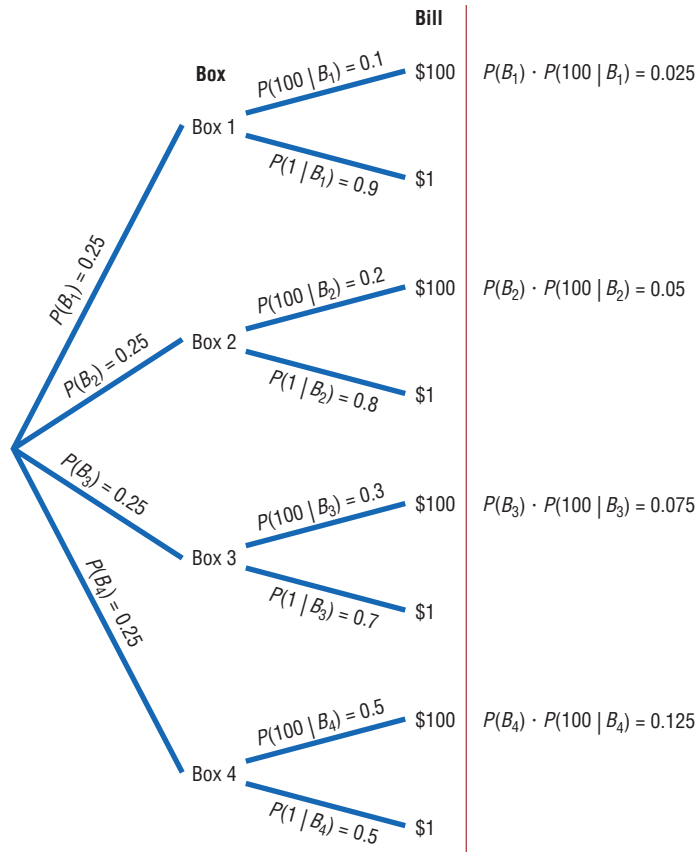
In Example B–2, the original probability of selecting box 4 was 0.25. However, once additional information was obtained—and the condition was considered that a $100 bill was selected—the revised probability of selecting box 4 became 0.455.

Bayes' theorem can be used to revise probabilities of events once additional information becomes known. Bayes' theorem is used as the basis for a branch of statistics called *Bayesian decision making,* which includes the use of subjective probabilities in making statistical inferences.

Exercises

B–1. An appliance store purchases electric ranges from two companies. From company A, 500 ranges are purchased and 2% are defective. From company B, 850 ranges are purchased and 2% are defective. Given that a range is defective, find the probability that it came from company B.

B–2. Two manufacturers supply blankets to emergency relief organizations. Manufacturer A supplies 3000 blankets, and 4% are irregular in workmanship. Manufacturer B supplies 2400 blankets, and 7% are found to be irregular. Given that a blanket is irregular, find the probability that it came from manufacturer B.

B–3. A test for a certain disease is found to be 95% accurate, meaning that it will correctly diagnose the disease in 95 out of 100 people who have the ailment. For a certain segment of the population, the incidence of the disease is 9%. If a person tests positive, find the probability that the person actually has the disease. The test is also 95% accurate for a negative result.

B–4. Using the test in Exercise B–3, if a person tests negative for the disease, find the probability that the person actually has the disease. Remember, 9% of the population has the disease.

B–5. A corporation has three methods of training employees. Because of time, space, and location, it sends 20% of its employees to location A, 35% to location B, and 45% to location C. Location A has an 80% success rate. That is, 80% of the employees who complete the course will pass the licensing exam. Location B has a 75% success rate, and location C has a 60% success rate. If a person has passed the exam, find the probability that the person went to location B.

B–6. In Exercise B–5, if a person failed the exam, find the probability that the person went to location C.

B–7. A store purchases baseball hats from three different manufacturers. In manufacturer A's box, there are 12 blue hats, 6 red hats, and 6 green hats. In manufacturer B's box, there are 10 blue hats, 10 red hats, and 4 green hats. In manufacturer C's box, there are 8 blue hats, 8 red hats, and 8 green hats. A box is selected at random, and a hat is selected at random from that box. If the hat is red, find the probability that it came from manufacturer A's box.

B–8. In Exercise B–7, if the hat selected is green, find the probability that it came from manufacturer B's box.

B–9. A driver has three ways to get from one city to another. There is an 80% probability of encountering a traffic jam on route 1, a 60% probability on route 2, and a 30% probability on route 3. Because of other factors, such as distance and speed limits, the driver uses route 1 fifty percent of the time and routes 2 and 3 each 25% of the time. If the driver calls the dispatcher to inform him that she is in a traffic jam, find the probability that she has selected route 1.

B–10. In Exercise B–9, if the driver did not encounter a traffic jam, find the probability that she selected route 3.

B–11. A store owner purchases telephones from two companies. From company A, 350 telephones are purchased and 2% are defective. From company B, 1050 telephones are purchased and 4% are defective. Given that a phone is defective, find the probability that it came from company B.

B–12. Two manufacturers supply food to a large cafeteria. Manufacturer A supplies 2400 cans of soup, and 3% are found to be dented. Manufacturer B supplies 3600 cans, and 1% are found to be dented. Given that a can of soup is dented, find the probability that it came from manufacturer B.

Appendix B–3

Alternate Approach to the Standard Normal Distribution

The following procedure may be used to replace the cumulative area to the left procedure shown in Section 6–1. This method determines areas from the mean where $z = 0$.

Finding Areas Under the Standard Normal Distribution

For the solution of problems using the standard normal distribution, a four-step procedure may be used with the use of the Procedure Table shown.

STEP 1 Sketch the normal curve and label.

STEP 2 Shade the area desired.

STEP 3 Find the figure that matches the shaded area from the following procedure table.

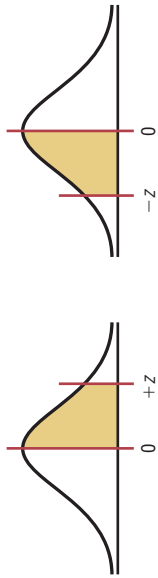**STEP 4** Follow the directions given in the appropriate block of the procedure table to get the desired area.

Note: Table B–1 gives the area between 0 and any z score to the right of 0, and all areas are positive.

There are seven basic types of problems and all seven are summarized in the Procedure Table, with appropriate examples.

Procedure Table
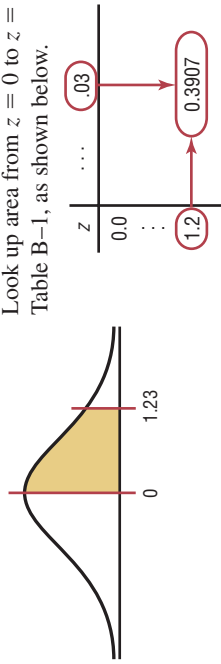
1. Between 0 and any z score:
Look up the z score in the table to get the area.

2. In any tail:
a. Look up the z score to get the area.
b. Subtract the area from 0.5000.

3. Between two z scores on the same side of the mean:
a. Look up both z scores to get the areas.
b. Subtract the smaller area from the larger area.

4. Between two z scores on opposite sides of the mean:
a. Look up both z scores to get the areas.
b. Add the areas.

Examples

B–3–1: Find the area between $z = 0$ and $z = 1.23$.

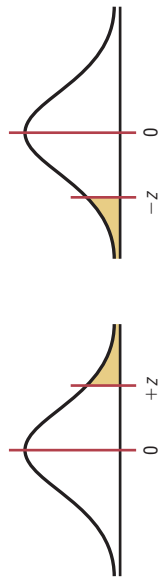Look up area from $z = 0$ to $z = 1.23$ on Table B–1, as shown below.

z	...	.03
0.0		
...		
1.2		0.3907

The area between $z = 0$ and $z = 1.23$ is 0.3907.

B–3–2: Find the area to the left of $z = -2.37$.

Look up area from $z = 0$ to $z = 2.37$ on Table B–1, as shown below.

z	...	.07
0.0		
...		
2.3		0.4911
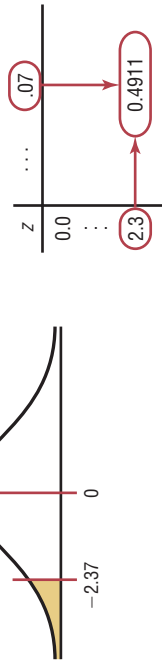
The area from $z = 0$ to $z = 2.37$ is the same as the area from $z = 0$ to $z = -2.37$.
Therefore, the area to the left of $z = -2.37 = 0.5000 - 0.4911 = 0.0089$.

B–3–3: Find the area between $z = 1.23$ and $z = 2.37$.

Look up areas for $z = 0$ to $z = 1.23$ and $z = 0$ to $z = 2.37$, as shown in B–3–1 and B–3–2, respectively.

The area between $z = 1.23$ and $z = 2.37 = 0.4911 - 0.3907 = 0.1004$.

B–3–4: Find the area between $z = -1.23$ and $z = 2.37$.

Look up areas for $z = 0$ to $z = 1.23$ and $z = 0$ to $z = 2.37$, as shown in B–3–1 and B–3–2, respectively.

The area between $z = -1.23$ and $z = 2.37 = 0.3907 + 0.4911 = 0.8818$.

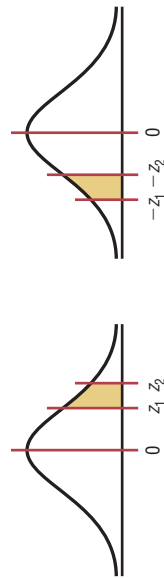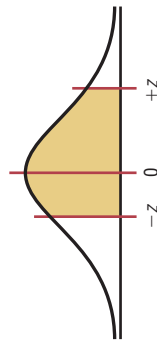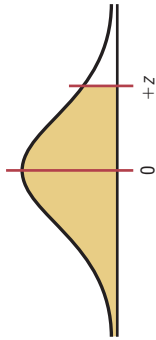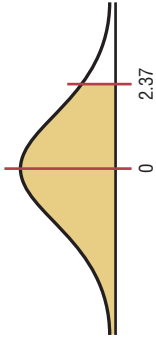5. To the left of any z score, where z is greater than the mean:
   a. Look up the z score to get the area.
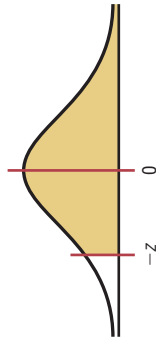   b. Add 0.5000 to the area.

B–3–5: Find the area to the left of $z = 2.37$.

Look up area for $z = 2.37$, as shown in B–3–2.
The area to the left of $z = 0$ is 0.5000.
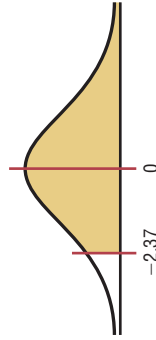The area to the left of $z = 2.37 = 0.4911 + 0.5000 = 0.9911$.

6. To the right of any z score, where z is less than the mean:
   a. Look up the area in the table to get the area.
   b. Add 0.5000 to the area.

B–3–6: Find the area to the right of $z = -2.37$.

Look up area for $z = 2.37$, as shown in B–3–2.
The area to the right of $z = 0$ is 0.5000.
The area to the right of $z = -2.37 = 0.4911 + 0.5000 = 0.9911$.

7. In any two tails:
   a. Look up the z scores in the table to get the areas.
   b. Subtract both areas from 0.5000.
   c. Add the answers.
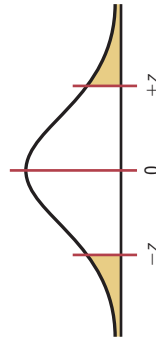
B–3–7: Find the area to the left of $z = -1.23$ and to the right of $z = 2.37$.

Look up areas for $z = 0$ to $z = 1.23$ and $z = 0$ to $z = 2.37$, as shown in B–3–1 and B–3–2, respectively.

Area to the left of $z = -1.23 = 0.5000 - 0.3907 = 0.1093$.

Area to the right of $z = 2.37 = 0.5000 - 0.4911 = 0.0089$.

The area to the left of $z = -1.23$ and to the right of $z = 2.37 = 0.1093 + 0.0089 = 0.1182$.

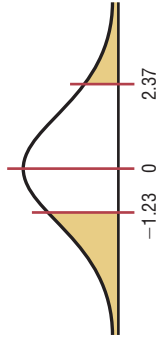| Table B–1	The Standard Normal Distribution								
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998

For z values greater than 3.49, use 0.4999.

Area given in table

0   z