CHAPTER 8

THE NORMAL DISTRIBUTION

OBJECTIVES

After completing this chapter, you should

- understand and be able to use the known characteristics of the normal curve to solve problems dealing with sample distributions in which you have to find areas under the curve and scores cutting off particular areas.

CHAPTER REVIEW

The normal distribution is important because many empirical distributions are similar to it and because it is the limiting case for a number of other important distributions in statistics. Although there are many possible normal curves, each *normal curve* is symmetrical and unimodal, and the tails of each curve never quite reach the baseline. The *standard normal curve* is a special example of the normal distribution in which the mean is 0 and the standard deviation is 1.

Almost the entire normal curve is contained within 6 standard deviation units or *z* scores. A *z score* is the deviation of a raw score from the mean in standard deviation units, and we use *z* scores to enter Table A, a table containing areas under the right half of the standard normal curve (see Appendix 2). In the chapter, Table A is used to answer questions about sample distributions, assuming the sample was drawn from a normally distributed population.

Two types of problems are described in the chapter: finding areas under the curve and finding a score or scores when an area or areas are known. For example, to find the percentile rank of a score, the raw score is converted to a *z* score, and Table A is consulted to determine the total area below the raw score. Similarly, to find the area above a score, the score is converted to a *z* score and Table A is consulted. To find an area between two scores, both scores are converted to *z* scores, and the appropriate areas from Table A are added together or the smaller is subtracted from the larger. *Percentage area* is the same thing as percentage frequency. Thus, if you are asked to give the number of subjects associated with an area under the curve, you first find the area and then take that percentage of the total *N*. Percentage area can also be converted to probability by dividing it by 100.

If an area is known, the *z* score associated with it can be found in Table A and converted into a raw score using Formula 6-19. One type of problem in which scores are to be found asks for deviant or unlikely scores without giving the direction of the deviance. In this case, the percentage area must be split in half and each half denoted in each tail of the normal curve. Two *z* scores are found from Table A, and both are converted to raw scores. If the problem gives the area as a probability, convert it to percentage area (multiply *p* by 100) before proceeding.

Before working any normal curve problem, it is helpful to draw a normal curve and label it with any information you have. Then use the curve to try to decide what you are looking for.

FORMULAS

The formulas for this chapter are two of the formulas covered at the end of Chapter 6: the formulas for converting any raw score to a *z* score and for converting any *z* score back to a raw score.

Formula 6-17. *Formula for finding a z score from a raw score using sample statistics*

$$z = \frac{X - \overline{X}}{s}$$

Formula 6-19. *Formula for finding a raw score from a z score using sample statistics*

$$X = zs + \overline{X}$$

TERMS TO DEFINE AND/OR IDENTIFY

normal distribution

bell curve

Gaussian distribution

probability distributions

standard normal curve

z score

percentile rank

FILL-IN-THE-BLANK ITEMS

Introduction

Because the normal distribution is useful in translating scores to probabilities, it is analogous to the (1) _________________ stone. Although the normal distribution is often attributed to German mathematician (2) _________________, it was actually introduced by (3) _________________. The normal distribution is important because many (4) _________________ distributions are similar to it. (5) _________________ distributions are those based on actual measurement. Also, the normal distribution is important for inferential statistics because it is the (6) _________________ case for a number of other important distributions. Examples of distributions that approach the normal distribution with large sample sizes are the chi-square distribution and the sampling distribution of (7) _________________.

Curves and Probability

All distributions of scores can be thought of as (8) _________________ distributions. The area under a portion of the curve represents the (9) _________________ associated with the scores falling in that area. Determining that probability usually involves two steps. First, we convert our raw scores to (10) _________________. Second, we use the *z* scores in conjunction with Table (11) _________________ in Appendix 2 to determine an (12) _________________ of the curve.

Characteristics of the Normal Curve

Each normal curve is (13) _________________; that is, the two halves coincide. Each normal curve has the same measures of (14) _________________ tendency, and the (15) _________________ of each curve never reach the baseline. The (16) _________________ normal curve has a mean of 0 and a standard deviation of 1. Almost the entire normal curve is bounded by (17) _________________ standard deviation units.

Review of *z Scores*

A (18) _________________ is the deviation of a raw score from the mean in standard deviation units. Negative *z* scores tell us that the raw score we are converting is (19) _________________ the mean, and positive values tell us that the raw score is (20) _________________ the mean.

Using the Normal Curve Table

In Table A, the percentage area between the mean and any *z* score is found in column (21) _________________. The remaining area beyond the *z* score is contained in column (22) _________________.

Finding Areas Under the Curve

Finding the percentile rank of a score

(23) _________________ _________________ is the percentage of cases up to and including the one in which we are interested. The first step in working any normal curve problem is to (24) _________________ the normal curve and label it with the information given in the problem. To find the percentile rank of a score, first convert the score to a (25) _________________. Then obtain the appropriate area from either column B or column C in Table A.

Finding the percentage of the normal curve above a score

After the normal curve is drawn and labeled appropriately, the (26) _________________ _________________ is converted to a *z* score. Then the appropriate area is obtained from Table (27) _________________. The *z* score itself is found in column (28) _________________.

Finding percentage frequency

When we find a percentage area under the normal curve, we can take that percentage of the total sample size to find (29) _________________ _________________ subjects have scores in the area.

Finding an area between two scores

To find an area between two scores, both scores must be converted to (30) _________________. Then either the areas from Table A are (31) _________________ or the smaller area is subtracted from the larger to find the area between the scores.

Probability and areas under the curve

The percentage areas under the curve can be converted into probabilities by dividing them by (32) _________________. The range of probability is from 0 to (33) _________________.

Finding Scores Cutting Off Areas

Finding the score that has a particular percentile rank

To locate a score associated with a particular area, first determine a (34) _________________ from Table A and then convert it to a raw score with Formula 6-19.

Finding deviant scores

When the problem asks for deviant or unlikely scores without specifying the direction of the deviance, you are really being asked to find scores at (35) _________________ ends of the distribution. The deviant percentage must first be divided in (36) _________________ before the graph can be correctly labeled.

Probability and deviant scores

If the problem of finding deviant scores is stated in terms of probability (e.g., find the scores that are so deviant that their probability is .05 or less), the first step is to convert the probability to (37) _________________ _________________. To do this, you multiply the probability by (38) _________________.

Troubleshooting Your Computations

As an aid to understanding and to determining whether your answer is appropriate, a small (39) _________________ _________________ should always be drawn and labeled as completely as possible. Then the obtained answer should be compared to the curve to see if it appears

(40) _________________. A common error is to not have the answer in the correct final

(41) _________________.

PROBLEMS

1. Assuming that a national certification examination for a professional group has a mean score of 72 and $s = 14.5$, answer the following questions.

a. What is the z score for a test score of 89.6?

b. What is the z score for a test score of 61.5?

c. What test score is 1.6 standard deviation units below the mean?

d. What test score is 1.6 standard deviation units above the mean?

e. What test scores are at least 1.5 standard deviation units away from the mean?

2. Use Table A in Appendix 2 to answer the following questions.

a. What is the percentage area between the mean and a z score of +0.57?

b. What is the percentage area between the mean and a z score of –0.57?

c. Between the mean and what z score lies approximately 46% of the normal curve? Can your answer be both positive and negative?

d. What z score cuts off the upper 10% of the distribution?

e. What area lies between a z score of +1.28 and a z score of +1.96?

f. What is the area above a z score of 1.96?

g. What is the total area between z scores of ±2.58?

h. What is the area below a z score of –1.28?

3. For 55 graduating seniors who took the Miller Analogies Test (MAT), the following results were obtained: $\Sigma X = 2{,}942.5$, $\Sigma X^2 = 165{,}606.68$. Use the normal distribution to answer the following questions.

a. What is the percentile rank of a score of 71?

b. What is the percentile rank of a score of 35?

c. How many seniors scored between 31 and 45?

d. What is the score at the 95th percentile?

e. What is the score at the 15th percentile?

f. A university has found that scores on the MAT are highly associated with success in its graduate program in English. Specifically, seniors scoring below 41 on the test almost always fail to earn degrees, and for this reason the university will not admit a person with this score or lower. How many of the 55 seniors should not bother to apply?

g. What MAT scores are so deviant that they occur 15% or less of the time?

h. What MAT scores are so unlikely that they occur with a probability of .01 or less?

4. A blood pressure testing machine is stationed in a busy corridor in a large shopping mall. For 1 week, a medical student records the diastolic pressure reading of each person who uses the machine between the hours of 10 A.M. and 2 P.M., with the following results: $N = 124$, $\Sigma X = 9{,}771.2$, $\Sigma X^2 = 828{,}360.92$.

a. How many of the sample had scores of 110 or higher?

b. What was the probability of a score of 60 or less?

c. If the normal range for diastolic pressures is between 60 and 90, how many of the 124 persons had a normal reading?

d. Based on these data, what percentage of the population had readings as deviant as 110?

e. What readings were so deviant that less than 5% of the sample had them?

5. A national survey of speeds on interstate highways found that the average speed of 9,549 automobiles was 67.3 mph, with a standard deviation of 7.81 mph.

a. What speed is at least 1.5 standard deviation units above the mean?

b. How many cars from the sample of 9,549 had speeds at least 1.5 standard deviation units above the mean?

c. What is the percentile rank of a speed of 55?

d. What is the percentile rank of a speed of 75?

e. How many automobiles had speeds 2 or more standard deviation units away from the mean?

f. What speeds were so deviant that their probability of occurrence was .05 or less?

CHECKING YOUR PROGRESS: A SELF-TEST

1. How does the standard normal curve differ from any other normal curve?

2. True or False: Areas under the normal curve below the mean are always negative.

3. Applicants for a job take a standardized test of their job-relevant skills. Assume that the scores of the 520 applicants are normally distributed with a mean of 48 and a standard deviation of 8.2.

   a. How many applicants scored higher than 65?

   b. How many applicants scored lower than 40?

c. What is the percentile rank of a score of 44?

d. What is the probability of a score of 60 or higher?

e. What score would an applicant have to obtain to be in the upper 10% of applicants?

f. What scores were so deviant that less than 2% of the sample had them?