Chapter 2 Chapter Review

This chapter focused on two big issues: describing an individual value's location within a distribution of data and modeling distributions with density curves. Both z-scores and percentiles provide easily calculated measures of position for individuals. Density curves come in assorted shapes, but all share the property that the area beneath the curve is 1. We can use areas under density curves to estimate the proportion of individuals in a distribution whose values fall in a specified range. In the special case of Normally distributed data, we can use the standard Normal curve and Table A to calculate such areas. There are many real-world examples of Normal distributions. When you meet a new set of data, you can use the graphical and numerical tools discussed in this chapter to assess the Normality of the data.

Chapter Review Exercises

These exercises are designed to help you review the important ideas and methods of the chapter. Relevant learning objectives are provided in bulleted form before each exercise.

- Find the standardized value (z-score) of an observation. Interpret z-scores in context.
- Use percentiles to locate individual values within distributions of data.

Exercises R2.1 and R2.2 refer to the following setting. According to the National Center for Health Statistics, the distribution of heights for 15-year-old males is symmetric, single-peaked, and bell-shaped. For this distribution, a z-score of 0 corresponds to a height of 170 centimeters (cm) and a z-score of 1 corresponds to a height of 177.5 cm.

R2.1 Male heights Consider the height distribution for 15-year-old males.

(a) Find its mean and standard deviation. Show your method clearly.

(b) What height would correspond to a z-score of 2.5? Show your work.

R2.2 Is Paul tall? Paul is 15 years old and 179 cm tall.

(a) Find the z-score corresponding to Paul's height. Explain what this value means.

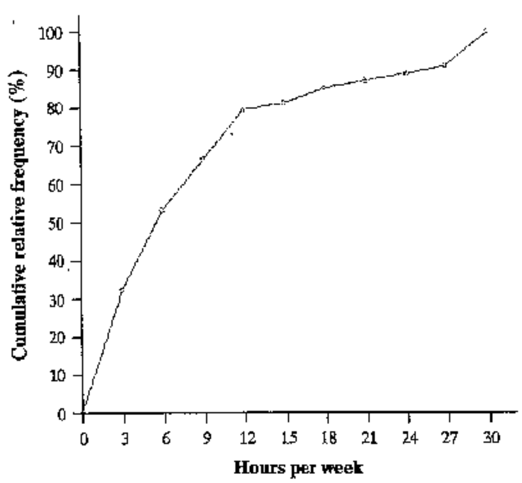
(b) Paul's height puts him at the 85th percentile among 15-year-old males. Explain what this means to someone who knows no statistics.

- Interpret a cumulative relative frequency graph.

R2.3 Computer use Mrs. Causey asked her students how much time they had spent using a computer during the previous week. The following figure shows a cumulative relative frequency graph of her students' responses.

(a) At what percentile does a student who used her computer for 10 hours last week fall?

(b) Estimate the median, Q_1, and Q_3 from the graph. Are there any outliers?

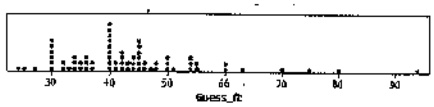


- Describe the effect of adding, subtracting, multiplying by, or dividing by a constant on the shape, center, and spread of a distribution of data.

R2.4 Aussie, Aussie, Aussie A group of Australian students were asked to estimate the width of their classroom in feet. Use the dotplot and summary statistics below to answer the following questions.

(a) Suppose we converted each student's guess from feet to meters (3.28 ft = 1 m). How would the shape of the distribution be affected? Find the mean, median, standard deviation, and *IQR* for the transformed data.

(b) The actual width of the room was 42.6 feet. Suppose we calculated the error in each student's guess as follows: guess − 42.6. Find the mean and standard deviation of the errors. How good were the students' guesses? Justify your answer.



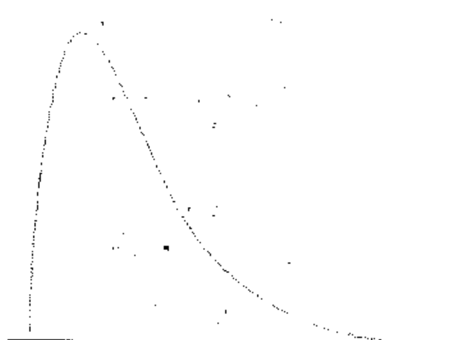
Variable	N	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Guess_ft	69	43.70	12.50	24.00	35.50	42.00	48.00	94.00

- Approximately locate the median (equal-areas point) and the mean (balance point) on a density curve.

R2.5 What the mean means The figure below is a density curve. Trace the curve onto your paper.

(a) Mark the approximate location of the median. Justify your choice of location.

(b) Mark the approximate location of the mean. Justify your choice of location.



- Use the 68–95–99.7 rule to estimate the percent of observations from a Normal distribution that fall in an interval involving points one, two, or three standard deviations on either side of the mean.

R2.6 Horse pregnancies Bigger animals tend to carry their young longer before birth. The length of horse pregnancies from conception to birth varies according to a roughly Normal distribution with mean 336 days and standard deviation 3 days. Use the 68–95–99.7 rule to answer the following questions.

(a) Almost all (99.7%) horse pregnancies fall in what range of lengths?

(b) What percent of horse pregnancies are longer than 339 days? Show your work.

- Use the standard Normal distribution to calculate the proportion of values in a specified interval.

R2.7 Standard Normal areas Use Table A to find the proportion of observations from a standard Normal distribution that falls in each of the following regions. In each case, sketch a standard Normal curve and shade the area representing the region.

(a) $z \leq -2.25$ (c) $z > 1.77$

(b) $z \geq -2.25$ (d) $-2.25 < z < 1.77$

- Use the standard Normal distribution to determine a *z*-score from a percentile.

R2.8 Working backward

(a) Find the number z at the 80th percentile of a standard Normal distribution.

(b) Find the number z such that 35% of all observations from a standard Normal distribution are greater than z.

- Use Table A to find the percentile of a value from any Normal distribution and the value that corresponds to a given percentile.

R2.9 Low-birth-weight babies Researchers in Norway analyzed data on the birth weights of 400,000 newborns over a 6-year period. The distribution of birth weights is Normal with a mean of 3668 grams and a standard deviation of 511 grams.[17] Babies that weigh less than 2500 grams at birth are classified as "low birth weight."

(a) What percent of babies will be identified as low birth weight? Show your work.

(b) Find the quartiles of the birth weight distribution. Show your work.

R2.10 Grading managers Many companies "grade on a bell curve" to compare the performance of their managers and professional workers. This forces the use of some low performance ratings, so that not all workers are listed as "above average." Ford Motor Company's "performance management process" for a time assigned 10% A grades, 80% B grades, and 10% C grades to the company's 18,000 managers. Suppose that Ford's performance scores really are Normally distributed. This year, managers with scores less than 25 received C's, and those with scores above 475 received A's. What are the mean and standard deviation of the scores? Show your work.

- Make an appropriate graph to determine if a distribution is bell-shaped.
- Use the 68–95–99.7 rule to assess Normality of a data set.

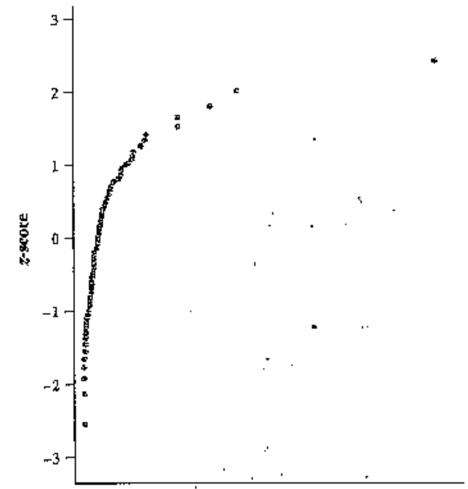
R2.11 Fruit fly thorax lengths Here are the lengths in millimeters of the thorax for 49 male fruit flies:[18]

0.64	0.64	0.64	0.68	0.68	0.68	0.72	0.72	0.72	0.72	0.74	0.76	0.76
0.76	0.76	0.76	0.76	0.76	0.76	0.78	0.80	0.80	0.80	0.80	0.80	0.82
0.82	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.88	0.88
0.88	0.88	0.88	0.88	0.88	0.88	0.92	0.92	0.92	0.94			

Are these data approximately Normally distributed? Give appropriate graphical and numerical evidence to support your answer.

- Interpret a Normal probability plot.

R2.12 Assessing Normality A Normal probability plot of a set of data is shown here. Would you say that these measurements are approximately Normally distributed? Why or why not?



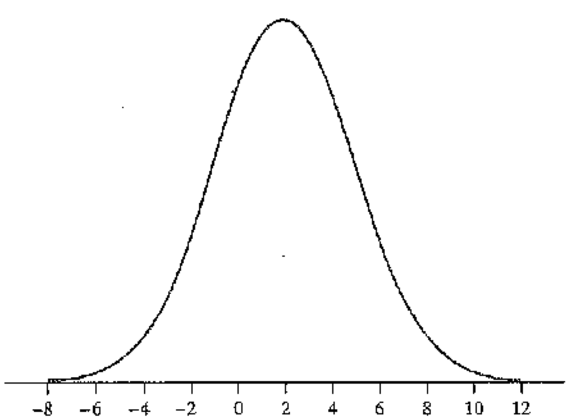
Chapter 2 AP Statistics Practice Test

Section I: Multiple Choice *Select the best answer for each question.*

T2.1. Many professional schools require applicants to take a standardized test. Suppose that 1000 students take such a test. Several weeks after the test, Pete receives his score report: he got a 63, which placed him at the 73rd percentile. This means that

(a) Pete's score was below the median.
(b) Pete did worse than about 63% of the test takers.
(c) Pete did worse than about 73% of the test takers.
(d) Pete did better than about 63% of the test takers.
(e) Pete did better than about 73% of the test takers.

T2.2. For the Normal distribution shown, the standard deviation is closest to

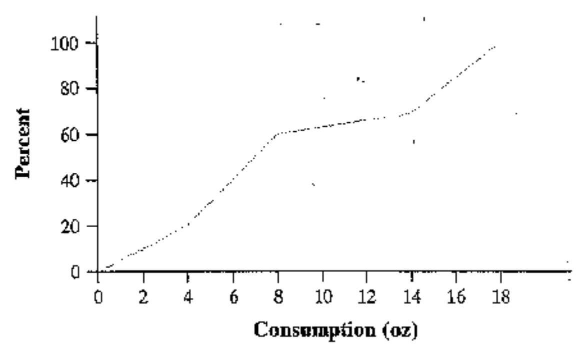


(a) 0 (b) 1 (c) 2 (d) 4 (e) 5

T2.3. Rainwater was collected in water collectors at 30 different sites near an industrial complex, and the amount of acidity (pH level) was measured. The mean and standard deviation of the values are 4.60 and 1.10, respectively. When the pH meter was recalibrated back at the laboratory, it was found to be in error. The error can be corrected by adding 0.1 pH units to all of the values and then multiplying the result by 1.2. The mean and standard deviation of the corrected pH measurements are

(a) 5.64, 1.44 (b) 5.64, 1.32 (c) 5.40, 1.44 (d) 5.40, 1.32 (e) 5.64, 1.20

T2.4. The figure shows a cumulative relative frequency graph of the number of ounces of alcohol consumed per week in a sample of 150 adults. About what percent of these adults consume between 4 and 8 ounces per week?

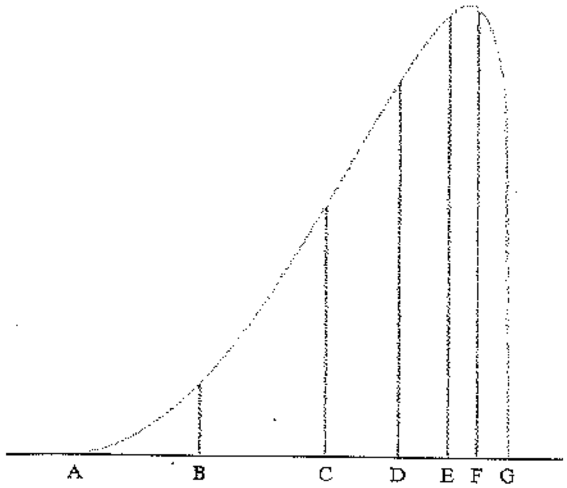


(a) 20% (b) 40% (c) 50% (d) 60% (e) 80%

T2.5. The average yearly snowfall in Chillyville is Normally distributed with a mean of 55 inches. If the snowfall in Chillyville exceeds 60 inches in 15% of the years, what is the standard deviation?

(a) 4.83 inches
(b) 5.18 inches
(c) 6.04 inches
(d) 8.93 inches
(e) The standard deviation cannot be computed from the given information.

T2.6. The figure shown is the density curve of a distribution. Five of the seven points marked on the density curve make up the five-number summary for this distribution. Which two points are *not* part of the five-number summary?



(a) B and E (c) C and E (e) A and G
(b) C and F (d) B and F

T2.7. If the heights of American men follow a Normal distribution, and 99.7% have heights between 5′0″ and 7′0″, what is your estimate of the standard deviation of the height of American men?

(a) 1″ (b) 3″ (c) 4″ (d) 6″ (e) 12″

T2.8. Which of the following is *not* correct about a standard Normal distribution?

(a) The proportion of scores that satisfy $0 < z < 1.5$ is 0.4332.

(b) The proportion of scores that satisfy $z < -1.0$ is 0.1587.

(c) The proportion of scores that satisfy $z > 2.0$ is 0.0228.

(d) The proportion of scores that satisfy $z < 1.5$ is 0.9332.

(e) The proportion of scores that satisfy $z > -3.0$ is 0.9938.

Questions T2.9 and T2.10 refer to the following setting. Until the scale was changed in 1995, SAT scores were based on a scale set many years ago. For Math scores, the mean under the old scale in the 1990s was 470 and the standard deviation was 110. In 2009, the mean was 515 and the standard deviation was 116.

T2.9. What is the standardized score (*z*-score) for a student who scored 500 on the old SAT scale?

(a) −30 (b) −0.27 (c) −0.13 (d) 0.13 (e) 0.27

T2.10. Jane took the SAT in 1994 and scored 500. Her sister Colleen took the SAT in 2009 and scored 530. Who did better on the exam, and how can you tell?

(a) Colleen—she scored 30 points higher than Jane.

(b) Colleen—her standardized score is higher than Jane's.

(c) Jane—her standardized score is higher than Colleen's.

(d) Jane—the standard deviation was bigger in 2009.

(e) The two sisters did equally well—their *z*-scores are the same.

Section II: Free Response

Show all your work. Indicate clearly the methods you use, because you will be graded on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

T2.11. As part of the President's Challenge, students can attempt to earn the Presidential Physical Fitness Award or the National Physical Fitness Award by meeting qualifying standards in five events: curl-ups, shuttle run, sit and reach, one-mile run, and pull-ups. The qualifying standards are based on the 1985 School Population Fitness Survey. For the Presidential award, the standard for each event is the 85th percentile of the results for a specific age group and gender among students who participated in the 1985 survey. For the National award, the standard is the 50th percentile. To win either award, a student must meet the qualifying standard for all five events.

Jane, who is 9 years old, did 40 curl-ups in one minute. Matt, who is 12 years old, also did 40 curl-ups in one minute. The qualifying standard for the Presidential award is 39 curl-ups for Jane and 50 curl-ups for Matt. For the National award, the standards are 30 and 40, respectively.

(a) Compare Jane's and Matt's performances using percentiles. Explain in language simple enough for someone who knows little statistics to understand.

(b) Who has the higher standardized value (*z*-score), Jane or Matt? Justify your answer.

T2.12. The army reports that the distribution of head circumference among male soldiers is approximately Normal with mean 22.8 inches and standard deviation 1.1 inches.

(a) A male soldier whose head circumference is 23.9 inches would be at what percentile? Show your method clearly.

(b) The army's helmet supplier regularly stocks helmets that fit male soldiers with head circumferences between 20 and 26 inches. Anyone with a head circumference outside that interval requires a customized helmet order. What percent of male soldiers require custom helmets? Show your work, including a well-labeled sketch of a Normal curve.

(c) Find the interquartile range for the distribution of head circumference among male soldiers. Show your method clearly.

T2.13. A study recorded the amount of oil recovered from the 64 wells in an oil field. Here are descriptive statistics for that set of data from Minitab.

Descriptive Statistics: Oilprod

Variable	N	Mean	Median	StDev	Min	Max	Q_1	Q_3
Oilprod	64	48.25	37.80	40.24	2.00	204.90	21.40	60.75

Does the amount of oil recovered from all wells in this field seem to follow a Normal distribution? Give appropriate statistical evidence to support your answer.