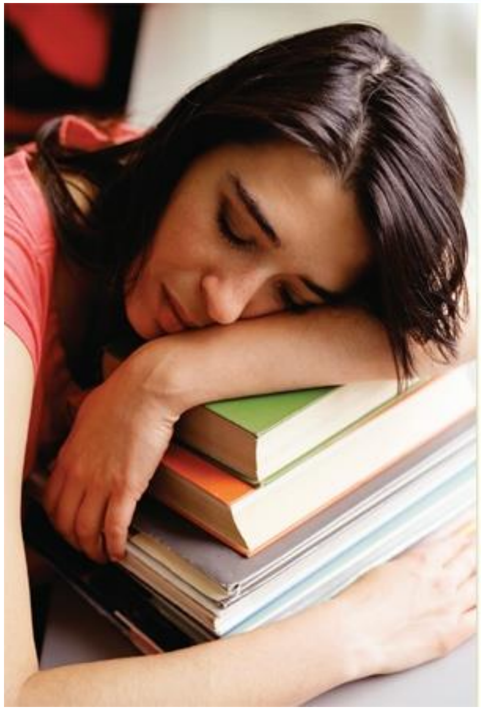
2 Descriptive Analysis and Presentation of Single-Variable Data



2.5 Measures of Position

Measures of Position

Measures of position are used to describe the position a specific data value possesses in relation to the rest of the data when in ranked order. *Quartiles* and *percentiles* are two of the most popular measures of position.

Quartiles Values of the variable that divide the ranked data into quarters; each set of data has three quartiles.

The *first quartile*, Q_1, is a number such that at most 25% of the data are smaller in value than Q_1 and at most 75% are larger. The *second quartile* is the median.

Measures of Position

The *third quartile*, Q_3, is a number such that at most 75% of the data are smaller in value than Q_3 and at most 25% are larger. (See Figure 2.24.)

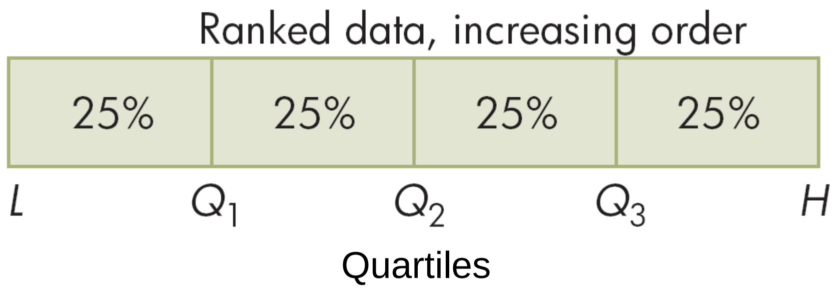


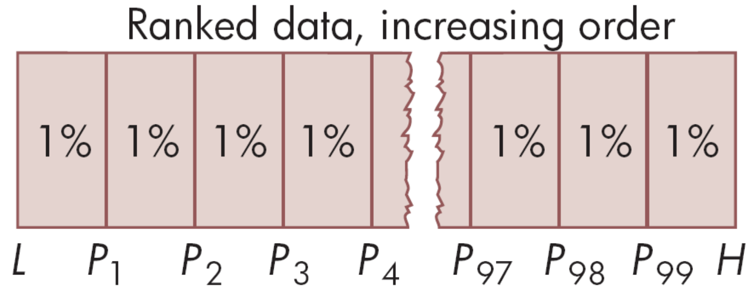
Figure 2.24

The procedure for determining the values of the quartiles is the same as that for percentiles and is shown in the following description of *percentiles*.

Remember that your data must be ranked from low (*L*) to high (*H*).

Measures of Position

Percentiles Values of the variable that divide a set of ranked data into 100 equal subsets; each set of data has 99 percentiles (see Figure 2.25).

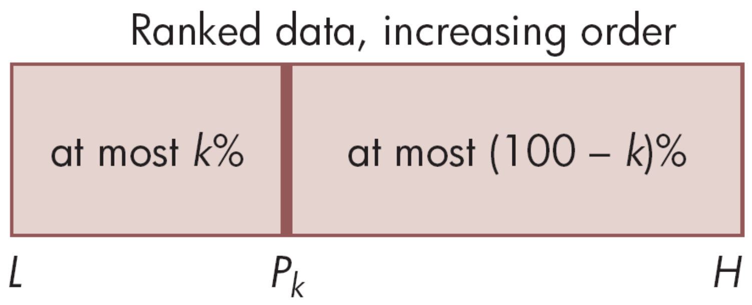


Percentiles

Figure 2.25

Measures of Position

The *k*th percentile, P_k, is a value such that at most *k*% of the data are smaller in value than P_k and at most (100 – *k*)% of the data are larger (see Figure 2.26).



*k*th Percentile

Figure 2.26

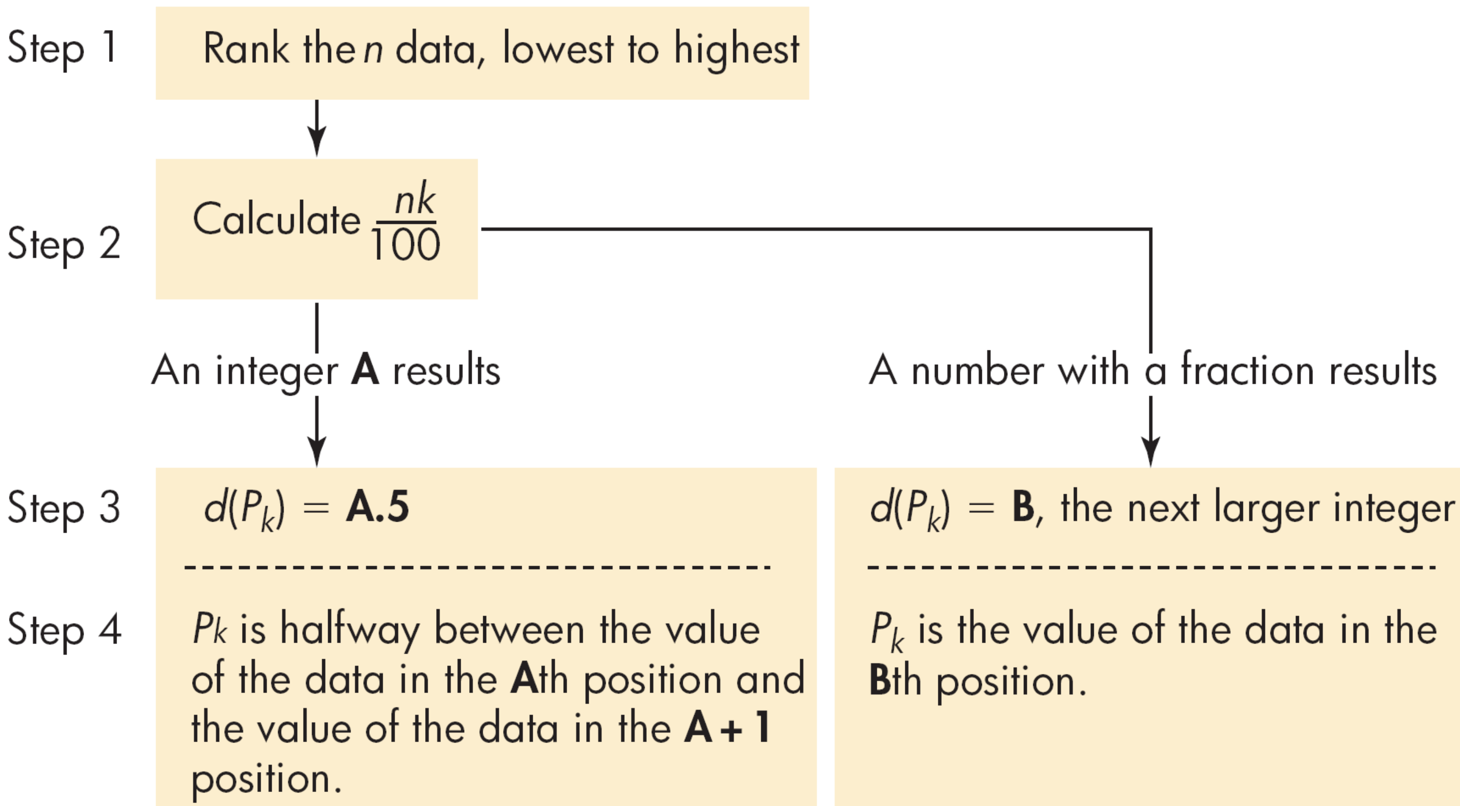
Measures of Position

Notes

1. The first quartile and the 25th percentile are the same; that is, $Q_1 = P_{25}$. Also, $Q_3 = P_{75}$.

2. The median, the second quartile, and the 50th percentile are all the same: $\tilde{x} = Q_2 = P_{50}$. Therefore, when asked to find P_{50} or Q_2, use the procedure for finding the median.

Measures of Position



Finding P_k Procedure

Figure 2.27

Example 12 – *Finding Quartiles and Percentiles*

Using the sample of 50 elementary statistics final exam scores listed in Table 2.15, find the first quartile, Q_1; the 58th percentile, P_{58}; and the third quartile, Q_3.

60	47	82	95	88	72	67	66	68	98	90	77	86
58	64	95	74	72	88	74	77	39	90	63	68	97
70	64	70	70	58	78	89	44	55	85	82	83	
72	77	72	86	50	94	92	80	91	75	76	78	

Raw Scores for Elementary Statistics Exam **[TA02-06]**

Table 2.15

Example 12 – *Solution*

Step 1: Rank the data: A ranked list may be formulated (see Table 2.16), or a graphic display showing the ranked data may be used.

39	64	72	78	89
44	66	72	80	90
47	$\boxed{67}$	74	82	90
50	68	74	82	91
55	68	75	83	92
58	70	76	85	94
58	70	77	86	95
60	70	77	$\boxed{86}$	95
63	72	$\boxed{77}$	88	97
64	72	$\boxed{78}$	88	98

13th position from *L*

29th and 30th positions from *L*

13th position from *H*

Stem-and-leaf of score *N* = 50

Leaf Unit = 1.0

1	3	9
2	4	4
3	4	7
4	5	0
7	5	5 8 8
11	6	0 3 4 4
15	6	6 $\boxed{7}$ 8 8
24	7	0 0 0 2 2 2 2 4 4
(7)	7	5 6 7 7 $\boxed{7\ 8}$ 8
19	8	0 2 2 3
15	8	5 6 $\boxed{6}$ 8 8 9
9	9	0 0 1 2 4
4	9	5 5 7 8

Ranked Data: Exam Scores

Table 2.16

Final Exam Scores

Figure 2.28

Example 12 – *Solution*

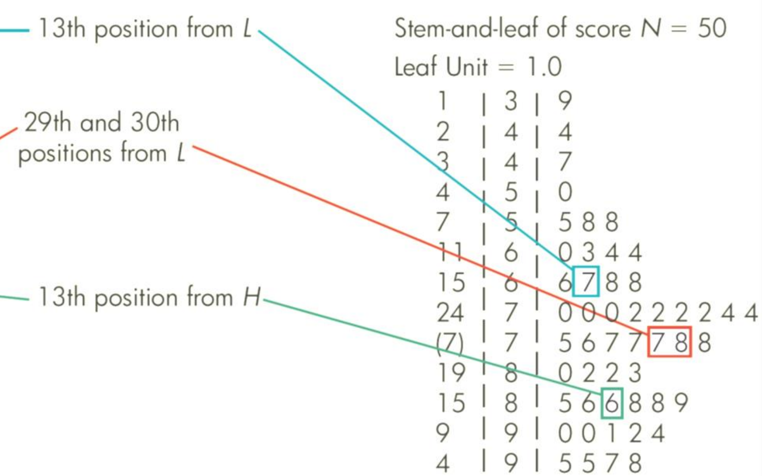
cont'd

The dotplot and the stem-and-leaf display are handy for this purpose. The stem-and-leaf display is especially helpful because it gives depth numbers counted from both extremes when it is computer generated (see Figure 2.28).

39	64	72	78	89
44	66	72	80	90
47	$\boxed{67}$	74	82	90
50	68	74	82	91
55	68	75	83	92
58	70	76	85	94
58	70	77	86	95
60	70	77	$\boxed{86}$	95
63	72	$\boxed{77}$	88	97
64	72	$\boxed{78}$	88	98

Ranked Data: Exam Scores

Table 2.16



Final Exam Scores

Figure 2.28

Example 12 – *Solution*

cont'd

Step 1 is the same for all three statistics.

Find Q_1:

Step 2: Find $\frac{nk}{100}$: $\frac{nk}{100} = \frac{(50)(25)}{100} = \mathbf{12.5}$

(n = 50 and k = 25, since $Q_1 = P_{25}$.)

Step 3: Find the depth of Q_1: $d(Q_1)$ = **13** (Since 12.5 contains a fraction, **B** is the next larger integer, 13.)

Step 4: Find Q_1: Q_1 is the 13th value, counting from L (see Table 2.16 or Figure 2.28), Q_1 = **67**

Example 12 – *Solution*

cont'd

Find P_{58}:

Step 2: Find $\frac{nk}{100}$: $\frac{nk}{100} = \frac{(50)(58)}{100} = \mathbf{29}$:

(n = 50 and k = 58 for P_{58}.)

Step 3: Find the depth of P_{58}: $d(P_{58})$ = **29.5** (Since **A** = 29, an integer, add 0.5 and use 29.5.)

Example 12 – *Solution*

cont'd

Step 4: Find P_{58}: P_{58} is the value halfway between the values of the 29th and the 30th pieces of data, counting from L (see Table 2.16 or Figure 2.28).

39	64	72	78	89
44	66	72	80	90
47	$\boxed{67}$	74	82	90
50	68	74	82	91
55	68	75	83	92
58	70	76	85	94
58	70	77	86	95
60	70	77	$\boxed{86}$	95
63	72	$\boxed{77}$	88	97
64	72	$\boxed{78}$	88	98

13th position from L

29th and 30th positions from L

13th position from H

Stem-and-leaf of score $N = 50$

Leaf Unit = 1.0

1	3	9
2	4	4
3	4	7
4	5	0
7	5	5 8 8
11	6	0 3 4 4
15	6	6 $\boxed{7}$ 8 8
24	7	0 0 0 2 2 2 2 4 4
(7)	7	5 6 7 7 $\boxed{7\ 8}$ 8
19	8	0 2 2 3
15	8	5 6 $\boxed{6}$ 8 8 9
9	9	0 0 1 2 4
4	9	5 5 7 8

Ranked Data: Exam Scores

Table 2.16

Final Exam Scores

Figure 2.28

Example 12 – *Solution*

cont'd

So

$$P_{58} = \frac{77 + 78}{2}$$

$$= 77.5$$

Therefore, it can be stated that "at most, 58% of the exam grades are smaller in value than 77.5." This is also equivalent to stating that "at most, 42% of the exam grades are larger in value than 77.5."

Optional technique: When k is greater than 50, subtract k from 100 and use $(100 – k)$ in place of k in Step 2. The depth is then counted from the highest-value data, H.

Example 12 – *Solution*

cont'd

Find Q_3 using the optional technique:

Step 2: Find $\frac{nk}{100}:\frac{nk}{100} = \frac{(50)(25)}{100} = \mathbf{12.5}$

(n = 50 and k = 75, since $Q_3 = P_{75}$, and k > 50; use $100 - k = 100 - 75 = 25$.)

Step 3: Find the depth of Q_3 from H: $d(Q_3) = 13$

Example 12 – *Solution*

cont'd

Step 4: Find Q_3: Q_3 is the 13th value; counting from *H* (see Table 2.16 or Figure 2.28), Q_3 = **86**

39	64	72	78	89
44	66	72	80	90
47	$\boxed{67}$	74	82	90
50	68	74	82	91
55	68	75	83	92
58	70	76	85	94
58	70	77	86	95
60	70	77	$\boxed{86}$	95
63	72	$\boxed{77}$	88	97
64	72	$\boxed{78}$	88	98

13th position from *L*

29th and 30th positions from *L*

13th position from *H*

Stem-and-leaf of score *N* = 50

Leaf Unit = 1.0

1	3	9
2	4	4
3	4	7
4	5	0
7	5	5 8 8
11	6	0 3 4 4
15	6	6 $\boxed{7}$ 8 8
24	7	0 0 0 2 2 2 2 4 4
(7)	7	5 6 7 7 $\boxed{7\ 8}$ 8
19	8	0 2 2 3
15	8	5 6 $\boxed{6}$ 8 8 9
9	9	0 0 1 2 4
4	9	5 5 7 8

Ranked Data: Exam Scores

Table 2.16

Final Exam Scores

Figure 2.28

Example 12 – *Solution*

cont'd

Therefore, it can be stated that "at most, 75% of the exam grades are smaller in value than 86." This is also equivalent to stating that "at most, 25% of the exam grades are larger in value than 86."

Measures of Position

Midquartile The numerical value midway between the first quartile and the third quartile.

$$\text{midquartile} = \frac{Q_1 + Q_3}{2} \qquad \textbf{(2.10)}$$

5-number summary The 5-number summary is composed of the following:

1. L, the smallest value in the data set
2. Q_1, the first quartile (also called P_{25}, the 25th percentile)
3. $\tilde{x}$, the median

Measures of Position

4. Q_3, the third quartile (also called P_{75}, the 75th percentile)
5. H, the largest value in the data set

Interquartile range The difference between the first and third quartiles. It is the range of the middle 50% of the data.

Box-and-whiskers display A graphic representation of the 5-number summary. The five numerical values (smallest, first quartile, median, third quartile, and largest) are located on a scale, either vertical or horizontal.

Measures of Position

The box is used to depict the middle half of the data that lie between the two quartiles.

The whiskers are line segments used to depict the other half of the data: One line segment represents the quarter of the data that are smaller in value than the first quartile, and a second line segment represents the quarter of the data that are larger in value than the third quartile.

Measures of Position

The position of a specific value can also be measured in terms of the mean and standard deviation using the *standard score*, commonly called the *z-score*.

Standard score, or *z*-score The position a particular value of *x* has relative to the mean, measured in standard deviations. The *z*-score is found by the formula

$$z = \frac{\text{value} - \text{mean}}{\text{st.dev.}} = \frac{x - \bar{x}}{s} \qquad \textbf{(2.11)}$$

Example 14 – *Finding z-Scores*

Find the standard scores for (a) 92 and (b) 72 with respect to a sample of exam grades that have a mean score of 74.92 and a standard deviation of 14.20.

Solution:

a. $x = 92$, $\bar{x} = 74.92$, $s = 14.20$. Thus,

$$z = \frac{x - \bar{x}}{s}$$

$$= \frac{92 - 74.92}{14.20}$$

$$= \frac{17.08}{14.20} = 1.20.$$

Example 14 – *Solution*

cont'd

b. $x = 72$, $\overline{x} = 74.92$, $s = 14.20$. Thus,

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

$$= \frac{72 - 74.92}{14.20}$$

$$= \frac{-2.92}{14.20} = -0.21.$$

This means that the score 92 is approximately 1.2 standard deviations above the mean and that the score 72 is approximately one-fifth of a standard deviation below the mean.

Measures of Position

Notes

1. Typically, the calculated value of z is rounded to the nearest hundredth.
2. z-scores typically range in value from approximately –3.00 to +3.00.

Because the z-score is a measure of relative position with respect to the mean, it can be used to help us compare two raw scores that come from separate populations.