STATISTICS

INFORMED DECISIONS USING DATA

Fifth Edition

Chapter 3

Numerically Summarizing Data

Pearson

3.3 Measures of Central Tendency and Dispersion from Grouped Data

Learning Objectives

1. Approximate the mean of a variable from grouped data
2. Compute the weighted mean → GPA
3. Approximate the standard deviation of a variable from grouped data

3.3 Measures of Central Tendency and Dispersion from Grouped Data

3.3.1 Approximate the Mean of a Variable from Grouped Data (1 of 4)

We have discussed how to compute descriptive statistics from raw data, but often the only available data have already been summarized in frequency distributions (**grouped data**).

Although we cannot find exact values of the mean or standard deviation without raw data, we can approximate these measures using the techniques discussed in this section.

3.3 Measures of Central Tendency and Dispersion from Grouped Data

3.3.1 Approximate the Mean of a Variable from Grouped Data (2 of 4)

Approximate the Mean of a Variable from a Frequency Distribution

is an approximation!

Population Mean

$$\mu = \frac{\sum x_i f_i}{\sum f_i} = \frac{x_1 f_1 + x_2 f_2 + \ldots + x_n f_n}{f_1 + f_2 + \ldots + f_n}$$

sum MP* freq.

sum freq = total

Sample Mean

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = \frac{x_1 f_1 + x_2 f_2 + \ldots + x_n f_n}{f_1 + f_2 + \ldots + f_n}$$

where x_i is the midpoint or value of the i^{th} class
f_i is the frequency of the i^{th} class
n is the number of classes

3.3 Measures of Central Tendency and Dispersion from Grouped Data

3.3.1 Approximate the Mean of a Variable from Grouped Data (3 of 4)

EXAMPLE Approximating the Mean from a Relative Frequency Distribution

The National Survey of Student Engagement is a survey that (among other things) asked first year students at liberal arts colleges how much time they spend preparing for class each week. The results from the 2007 survey are summarized below. Approximate the mean number of hours spent preparing for class each week.

Hours	0	1–5	6–10	11–15	16–20	21–25	26–30	31–35
Frequency	0	130	250	230	180	100	60	50

Source:
http://nsse.iub.edu/NSSE_2007_Annual_Report/docs/withhold/NSSE_2007_Annual_Report.pdf

3.3 Measures of Central Tendency and Dispersion from Grouped Data

3.3.1 Approximate the Mean of a Variable from Grouped Data (4 of 4)

Calc Shortcut:
Lists (L1, L2, L3,…)

L1 L2 L3= L1*L2

Time	Frequency	x_i	$x_i f_i$
0	0		
1 - 5	130	3	
6 - 10	250	8	
11 - 15	230	13	
16 - 20	180	18	
21 - 25	100	23	
26 - 30	60	28	
31 - 35	50	33	
	$\sum f_i = 1000$		$\sum x_i f_i =$ 14250

Calc Shortcut:

- "List" (2nd+Stat)
- Math
- Sum(
- Enter list (L3)

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i}$$ sum(L3) / sum(L1)

$= \frac{14250}{1000}$

$=$ 14.25 hours per student

3.3 Measures of Central Tendency and Dispersion from Grouped Data

3.3.1 Approximate the Mean of a Variable from Grouped Data (4 of 4)

Calc Shortcut:
Lists (L1, L2, L3,…)

L1 L2 L3= L1*L2

Time	Frequency	x_i	$x_i f_i$
0	0	0	0
1 - 5	130	3.5	455
6 - 10	250	8.5	2125
11 - 15	230	13.5	3105
16 - 20	180	18.5	3330
21 - 25	100	23.5	2350
26 - 30	60	28.5	1710
31 - 35	50	33.5	1675
	$\sum f_i = 1000$		$\sum x_i f_i = 14{,}750$

Calc Shortcut:

- "List" (2^{nd}+Stat)
- Math
- Sum(
- Enter list (L3)

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = \frac{14{,}750}{1000} = 14.75$$

3.3 Measures of Central Tendency and Dispersion from Grouped Data

3.3.2 Compute the Weighted Mean (1 of 2)

The **weighted mean**, $\overline{x}_w$, of a variable is found by multiplying each value of the variable by its corresponding weight, adding these products, and dividing this sum by the sum of the weights. It can be expressed using the formula

$$\overline{x}_w = \frac{\sum w_i x_i}{w_i} = \frac{w_1 x_1 + w_2 x_2 + \ldots + w_n x_n}{w_1 + w_2 + \ldots + w_n}$$

where w is the weight of the i^{th} observation x_i is the value of the i^{th} observation

GTQs: 1) GPA 2) average costs of different items

3.3 Measures of Central Tendency and Dispersion from Grouped Data

3.3.2 Compute the Weighted Mean (2 of 2)

EXAMPLE Computed a Weighted Mean

Bob goes to the "Buy the Weight" Nut store and creates his own bridge mix. He combines 1 pound of raisins, 2 pounds of chocolate covered peanuts, and 1.5 pounds of cashews. The raisins cost $1.25 per pound, the chocolate covered peanuts cost $3.25 per pound, and the cashews cost $5.40 per pound. What is the cost per pound of this mix?

3.3 Measures of Central Tendency and Dispersion from Grouped Data

3.3.2 Compute the Weighted Mean (2 of 2)

EXAMPLE Computed a Weighted Mean

Bob goes to the "Buy the Weigh" Nut store and creates his own bridge mix. He combines 1 pound of raisins, 2 pounds of chocolate covered peanuts, and 1.5 pounds of cashews. The raisins cost $1.25 per pound, the chocolate covered peanuts cost $3.25 per pound, and the cashews cost $5.40 per pound. What is the cost per pound of this mix?

$$\bar{x}_w = \frac{1(\$1.25) + 2(\$3.25) + 1.5(\$5.40)}{1 + 2 + 1.5}$$

$$= \frac{\$15.85}{4.5} = \$3.52$$

3.3 Measures of Central Tendency and Dispersion from Grouped Data

3.3.2 Compute the Weighted Mean (2 of 2)

EXAMPLE GPA

Patty has just completed her second semester in college. She earned a grade of A in her 2-hour discrete math course, a grade of D in her 5-hour economics course, a grade of D in her 2-hour engineering course, and a grade of C in her 1-hour creative writing course. Assuming that A equals 4 points, B equals 3 points, C equals 2 points, D equals 1 point, and F is worth no points, determine Patty's grade-point average for the semester.

Grade	Grade Point
A	4.0
B	3.0
C	2.0
D	1.0
F	0.0

	w	Grade	GP (x)
Discrete Math	2 hr	A	4.0
Econ	5 hr	D	1.0
Eng.	2 hr	D	1.0
Creative Write	1 hr	C	2.0

$$GPA = \frac{\sum w \cdot x}{\sum w} = \frac{2(4.0) + 5(1.0) + 2(1.0) + 1(2.0)}{2+5+2+1} = \frac{17.0}{10}$$

$$= 1.7 \text{ GPA}$$

3.3 Measures of Central Tendency and Dispersion from Grouped Data

3.3.2 Compute the Weighted Mean (2 of 2)

EXAMPLE GPA

Patty has just completed her second semester in college. She earned a grade of A in her 2-hour discrete math course, a grade of D in her 5-hour economics course, a grade of D in her 2-hour engineering course, and a grade of C in her 1-hour creative writing course. Assuming that A equals 4 points, B equals 3 points, C equals 2 points, D equals 1 point, and F is worth no points, determine Patty's grade-point average for the semester.

Grade | Grade Point

A | 4.0

B | 3.0

C | 2.0

D | 1.0

F | 0.0

$$\bar{x}_w = \frac{\sum w_i x_i}{w_i}$$

$$\bar{x}_w = \frac{2(4.0)+5(1.0)+2(1.0)+1(3.0)}{2+5+2+1} = \frac{18.0}{10} = 1.8 \text{ GPA}$$

Patty's grade point average in her second semester in college is a 1.8.