

Chapter 2

Describing Variables

2.7 Standardized Scores (Z Scores)

Mean and Standard Deviation

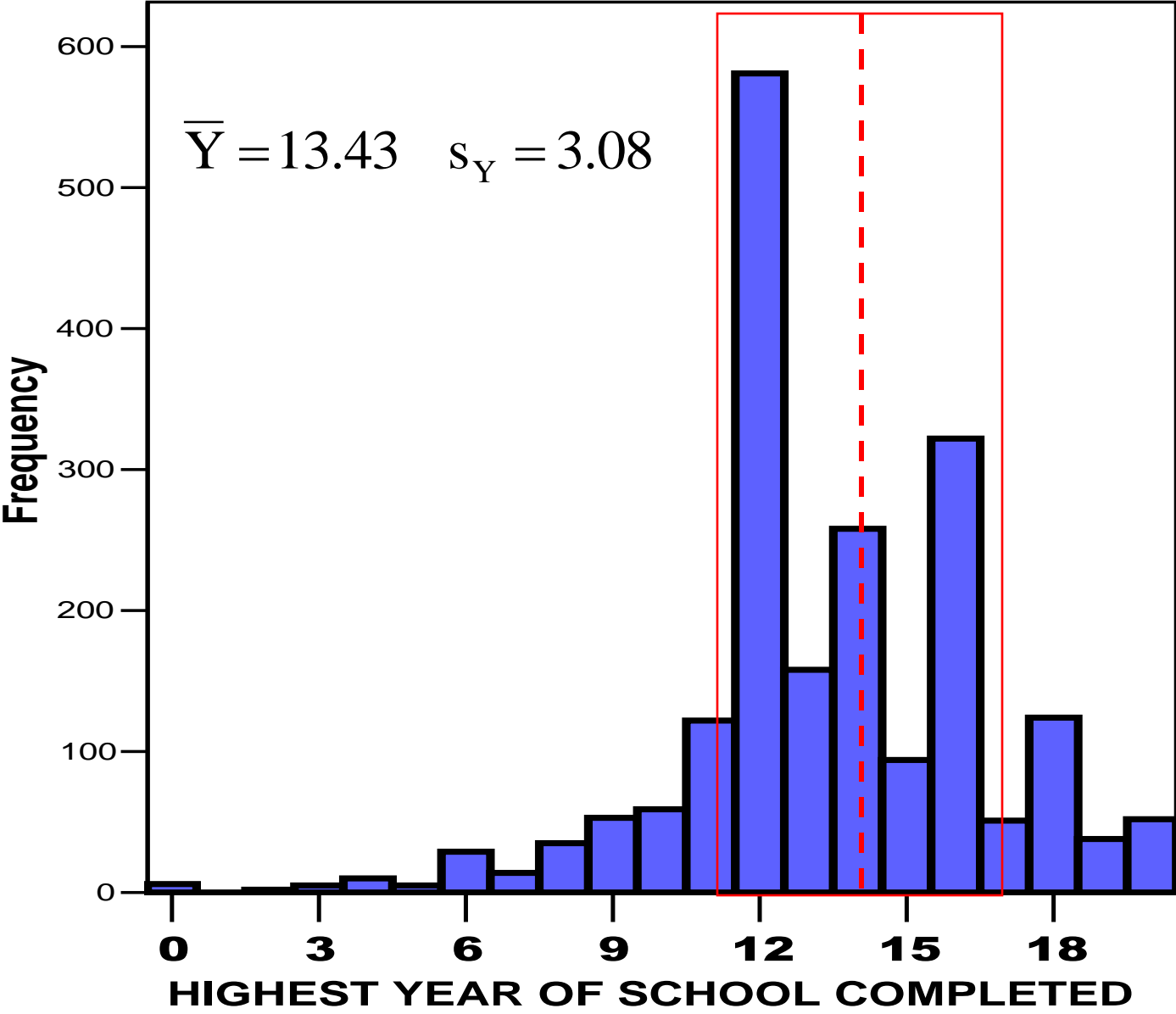
Two basic descriptive statistics for the distribution of a continuous variable with N observations:

Mean (central tendency): $\bar{Y} = \sum_{i=1}^N \frac{Y_i}{N}$

Standard deviation (dispersion):

$$s_Y = \sqrt{s_Y^2} = \sqrt{\sum_{i=1}^N \frac{(Y_i - \bar{Y})^2}{N-1}} = \sqrt{\frac{\sum d_i^2}{N-1}}$$

EDUC Years of Schooling



Standardized Scores (Z Scores)

Transform each score in the frequency distribution of a continuous Y -variable into a standard or Z score:

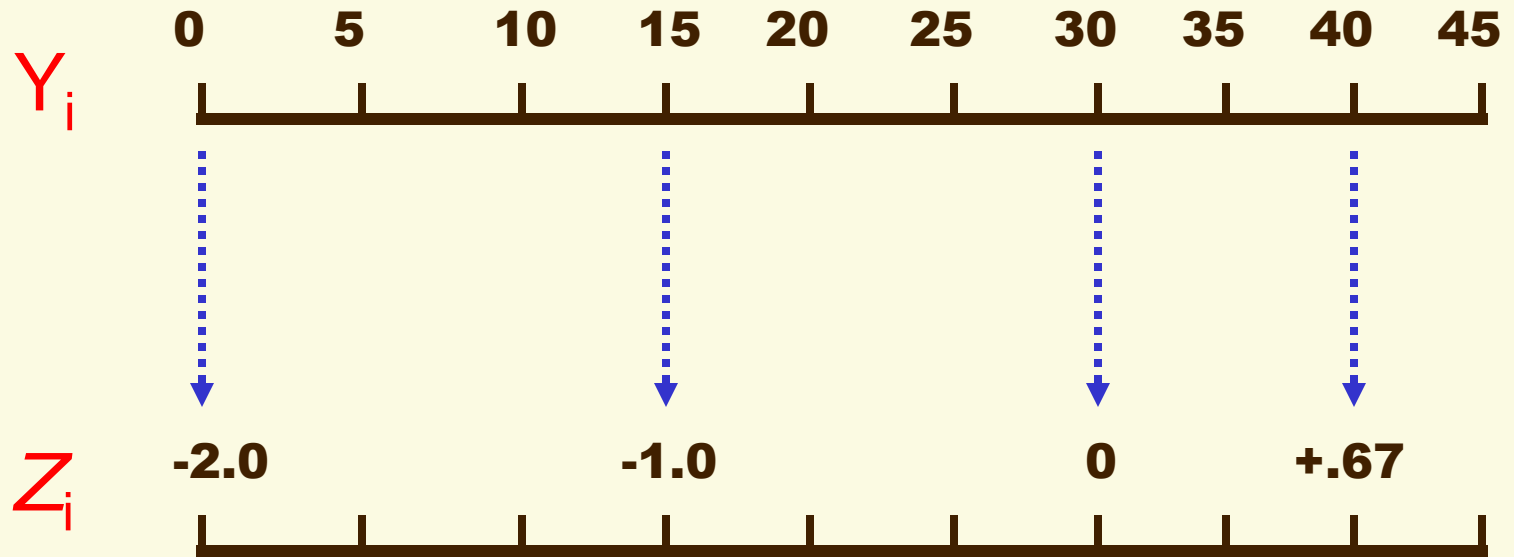
(1) from Y_i subtract the mean (= deviation d_i)

(2) divide case i deviation by the standard deviation

$$Z_i = \frac{Y_i - \bar{Y}}{S_Y} = \frac{d_i}{S_Y}$$

Z scores can be positive or negative, indicating the number of standard deviation units that Y_i lies above or below the distribution mean

Y_i scores into Z_i scores



$$Z_i = \frac{Y_i - \bar{Y}}{S_Y} \quad \bar{Y} = 30 \quad S_Y = 15$$

Change some EDUC Y scores into Z scores:

$$\bar{Y} = 13.43 \quad s_Y = 3.08$$

1. What is the Z score for a person with 16 years?

$$Z_i = \frac{Y_i - \bar{Y}}{s_Y} = \underline{\hspace{10cm}}$$

2. Find the Z score for a person with 8 years:

$$Z_i = \frac{Y_i - \bar{Y}}{s_Y} = \underline{\hspace{10cm}}$$

Now change some EDUC Z scores into Y scores:

What years of EDUC separate two persons at the upper & lower limits of the **range** $Z = \pm 2.2$ standard deviations?

- Find upper limit of range, $Z = +2.20$ std. dev. units :

$$Z_i = \frac{Y_i - \bar{Y}}{s_Y} \Rightarrow Y_i = +Z_i s_Y + \bar{Y}$$

$$Y_i = (+2.20)(3.08) + 13.43 = \underline{\hspace{2cm}}$$

- Find lower limit of range, $Z = -2.20$ std. dev. units :

$$Z_i = \frac{Y_i - \bar{Y}}{s_Y} \Rightarrow Y_i = -Z_i s_Y + \bar{Y}$$

$$Y_i = (-2.20)(3.08) + 13.43 = \underline{\hspace{2cm}}$$

Find the Z scores for these ungrouped data:

First calculate mean & standard deviation, then the Z scores

$$Y_i - \bar{Y} = d_i \quad (d_i)^2$$

Y₁: 6 - _____

Y₂: 4 - _____

Y₃: 5 - _____

Y₄: 4 - _____

Y₅: 3 - _____

Y₆: 8 - _____

$$\sum_{i=1}^6 (d_i)^2 = \underline{\hspace{2cm}}$$

$$s_Y^2 = \frac{\sum (d_i)^2}{N-1} = \underline{\hspace{2cm}}$$

$$s_Y = \sqrt{s_Y^2} = \underline{\hspace{2cm}}$$

$$Z_{Y=3} = \frac{Y_i - \bar{Y}}{s_Y} = \underline{\hspace{2cm}}$$

$$Z_{Y=6} = \underline{\hspace{2cm}}$$

$$Z_{Y=4} = \underline{\hspace{2cm}}$$

$$Z_{Y=8} = \underline{\hspace{2cm}}$$

Find Z scores for grouped data on Pres. Obama's handling of the global war on terror, where Mean = 3.30 and N = 158

Rating	Y_i	f_i	$(d_i)^2(f_i)$
Poor	1	10	_____
Fair	2	20	_____
Good	3	40	_____
Excellent	4	88	_____

$$Z_1 = (Y_1 - \bar{Y})/s_Y = \underline{\hspace{2cm}} \quad \sum_{i=1}^K (d_i)^2(f_i) = \underline{\hspace{2cm}}$$

$$Z_2 = (Y_2 - \bar{Y})/s_Y = \underline{\hspace{2cm}} \quad s_Y^2 = \frac{\sum_{i=1}^K (d_i)^2(f_i)}{N-1} = \underline{\hspace{2cm}}$$

$$Z_3 = (Y_3 - \bar{Y})/s_Y = \underline{\hspace{2cm}}$$

$$Z_4 = (Y_4 - \bar{Y})/s_Y = \underline{\hspace{2cm}} \quad s_Y = \underline{\hspace{2cm}}$$

Calculate the mean for ungrouped data

i	Y_i	
1	2	$\bar{Y} = \sum_{i=1}^N \frac{Y_i}{N}$
2	3	
3	3	
4	4	N = _____
5	5	$\sum_{i=1}^N Y_i = \text{_____}$
6	5	
7	6	
8	7	$\bar{Y} = \text{_____}$
9	7	
10	8	

Calculate variance & std. dev. for 10 scores

Y_i	-	\bar{Y}	=	d_i	$(d_i)^2$	
2	-	=	_____	_____		$\sum_{i=1}^{10} (d_i)^2 =$ _____
3	-	=	_____	_____		
3	-	=	_____	_____		$s_Y^2 = \sum (d_i)^2 / (N-1) =$
4	-	=	_____	_____		
5	-	=	_____	_____		_____
5	-	=	_____	_____		
6	-	=	_____	_____		$s_Y = \sqrt{s_Y^2} =$ _____
7	-	=	_____	_____		
7	-	=	_____	_____		
8	-	=	_____	_____		

Find All the Z Scores

$$Z_i = \frac{Y_i - \bar{Y}}{s_Y}$$

$(Y_i - \bar{Y})$	$/$	s_Y	$=$	
(2 -)	/	_____	=	_____
(3 -)	/	_____	=	_____
(3 -)	/	_____	=	_____
(4 -)	/	_____	=	_____
(5 -)	/	_____	=	_____
(5 -)	/	_____	=	_____
(6 -)	/	_____	=	_____
(7 -)	/	_____	=	_____
(7 -)	/	_____	=	_____
(8 -)	/	_____	=	_____