

## Measures of Dispersion (Spread)

1. **Range**—difference between largest and smallest values--unstable

$$R = x_L - x_S$$

2. **Percentiles or quantiles**

$p^{\text{th}}$  percentile is the value  $V_p$  such that  $p$  percent of the sample points are  $\leq V_p$ .

$P_{50}$  = median

**Def:** After sorting observations from smallest to largest,  $p^{\text{th}}$  percentile =

- a.  $(k+1)^{\text{th}}$  largest sample point if  $np/100$  is not an integer ( $k$  = largest integer less than  $np/100$ )
- b. average of the  $(np/100)^{\text{th}}$  and  $(np/100 + 1)^{\text{th}}$  largest observations if  $np/100$  is an integer

**Example:** For ER data, find the 25<sup>th</sup> percentile.

$n = 30, p = 25 \Rightarrow np/100 = 30*25/100 = 7.5$   
(not an integer)

Therefore, 25<sup>th</sup> percentile  
=  $(7 + 1)^{\text{th}}$  largest value = 34

Find the 50<sup>th</sup> percentile.

$n = 30, p = 50 \Rightarrow np/100 = 30*50/100 = 15$  (an integer)

Therefore, 50<sup>th</sup> percentile  
= average of  $(30*50/100)^{\text{th}}$  and  $(30*50/100 + 1)^{\text{th}}$  largest values

= average of 15<sup>th</sup> and 16<sup>th</sup> largest values

= average of 39, 43 = 41

Find  $P_{40}$ .

### Some Measures of Spread

$$P_{90} - P_{10}$$

$P_{75} - P_{25}$  – Called Interquartile Range

### 3. Variance

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

**Example: Fasting blood glucose levels of six children**

**74**  
**68**  
**60**  
**65**  
**71**  
**82**

$$\bar{x} = 70$$

$x_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
<b>74</b>	<b><math>74 - 70 = 4</math></b>	<b>16</b>
<b>68</b>	<b>-2</b>	<b>4</b>
<b>60</b>	<b>-10</b>	<b>100</b>
<b>65</b>	<b>-5</b>	<b>25</b>
<b>71</b>	<b>1</b>	<b>1</b>
<b>82</b>	<b>12</b>	<b>144</b>
<b>Total</b>	<b>0</b>	<b><math>290 = \sum (x_i - \bar{x})^2</math></b>

$$s^2 = 290 / 5 = 58$$

**Use an alternative formula if the data set is large or if the mean has a long decimal.**