Lecture 2.4B Standard Deviation from a Frequency Table ; Empirical Rule ; & Chebychev's Theorem

Assignment: Tues. p 90-97 9 & 10 : 33-42 all 46-50 evens

Standard Deviation from a Frequency Table:

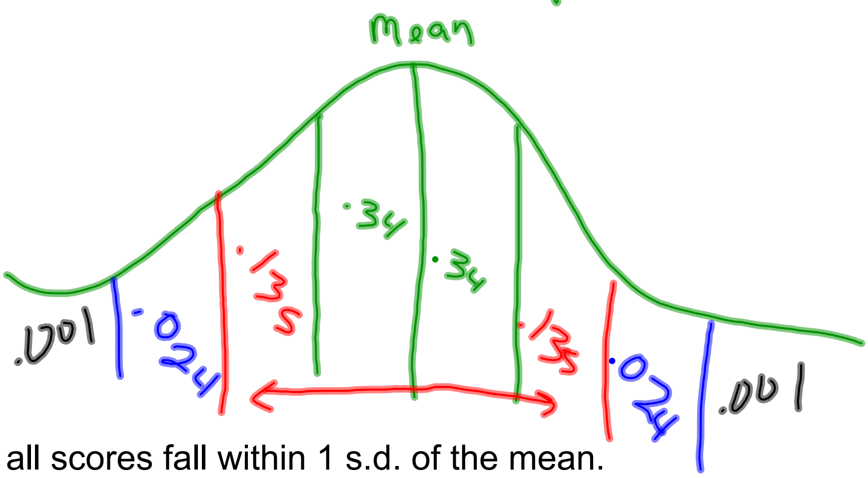
$$S = \sqrt{\frac{\Sigma (x - \bar{x})^2 \cdot f}{n - 1}}$$

class limits	frequencies	midpoints	$f \cdot x$	$x - \bar{x}$	$(x - \bar{x})^2$	$f \cdot (x - \bar{x})^2$
16 - 25	22	20.5	451	-15	225	4950
26 - 35	10	30.5	305	-5	25	250
36 - 45	6	40.5	243	5	25	150
46 - 55	2	50.5	101	15	225	450
56 - 65	4	60.5	242	25	625	2500
66 - 75	5	70.5	352.5	35	1225	6125
76 - 85	+ 1	80.5	80.5	45	2025	+ 2025
	50		1775 ÷ 50			16450

$\bar{x} = 35.5$

$S = \sqrt{\frac{16450}{50-1}} = 18.32$

Empirical Rule for the bell-shaped curve. p86

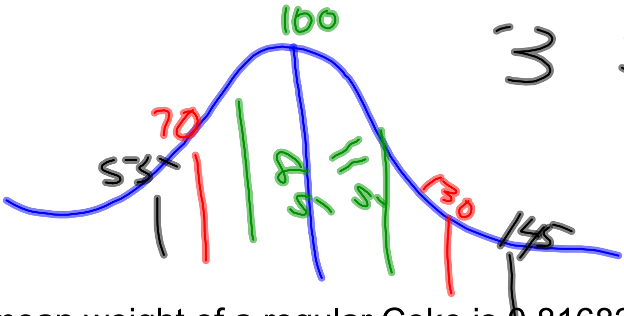


68% of all scores fall within 1 s.d. of the mean.

95% of all scores fall within 2 s.d. of the mean.

99.7 % of all scores fall within 3 s.d. of the mean.

Ex #1. Adult IQ scores have a bell-shaped distribution with a mean of 100 and a standard deviation of 15. Use the Empirical Rule to find the percentages of adults with IQ scores between 55 and 145.



3 S.D.

99.7%

Ex #2. The mean weight of a regular Coke is 0.81682 lbs., the standard deviation is 0.00751 lbs, and the distribution is bell-shaped. What percentage of cans of regular Coke will have weights between 0.80180 and 0.83184?

$$\begin{array}{r} .81682 \\ +\ .00751 \\ \hline .82433 \\ +\ .00751 \\ \hline .83184 \end{array}$$

2 SD

95%

Chebychev's Theorem is for any set of data (population or sample) and for any constant *k* greater than 1, the proportion of that data that must lie within *k* standard deviations on either side of the mean is a least $1 - \frac{1}{k^2}$.

Chebychev's Theorem applies to any and all distributions of data.

If the number of standard deviations = 2, then *k* = 2. So substitute 2 for *k* in the formula. $1 - \frac{1}{2^2} = 1 - \frac{1}{4} = \frac{3}{4}$ In other words, at least 3/4 or 75% of the data values lie within the range of 2 standard deviations below the mean to 2 standard deviations above the mean.

So we can figure out the percents for the number of standard deviations.

If $k = 3$, $1 - \frac{1}{3^2} = 1 - 1/9 = 8/9$ or 88.9%. So 88.9% of the scores lie within 3 standard deviations of the mean.

So if *k* = 4 , 94 % if *k* = 5, 96 % if *k* = 10 , 99 %

Ex #1. A data set has a mean of 25.2 hours and a standard deviation of 1.6 hours. Find an interval in which 93.8% of the data values will fall.

25.2 + 4(1.6) = 31.6 hrs to

25.2 − 4(1.6) = 18.8 hrs

Ex #2. The mean sale per customer for 40 customers at a grocery store is $23.00 with a standard deviation of $6.00. Using Chebychev's Theorem, at least how many of the customers spent between $11.00 and $35.00?

2 S.D = 75%

40 x .75 =

30 customers