Section 6.3:

Monday, March 18, 2024 11:01 AM

Tchebysheff's Theorem and the Empirical Rule

or Chebyshev

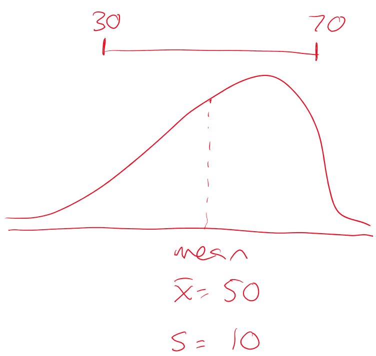
Tchebysheff's Theorem: works for <u>all</u> data sets

(symmetrical or skewed, unimodal / bimodal / mult)

- for <u>any</u> set of measurements

<u>at</u> <u>least</u> $\left(1 - \frac{1}{k^2}\right)$ of the measurements fall within k standard deviations of the mean for $k > 1$

example:



look at interval from 30 to 70

30 is two standard deviations below the mean, while 70 is two standard deviations above the mean

so $k = 2$ (two standard deviations away from the mean)

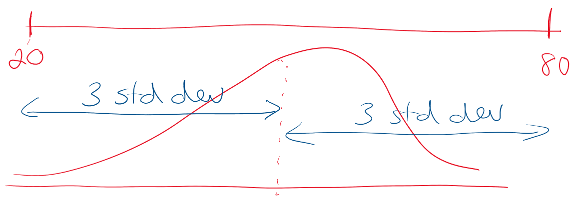
$1 - \frac{1}{k^2} = 1 - \frac{1}{2^2} = \frac{3}{4}$ or 75%

k^2 2^2 4

Conclusion: Tcheby says that at least 75% of the data fall between 30 and 70

Section 6.3: cont'd 2024/03/19

What about the number of measurements between 20 and 80?



$\bar{x} = 50$

$s = 10$

so $k = 3$

and $1 - \frac{1}{k^2} = 1 - \frac{1}{3^2} = \frac{8}{9}$ or 89%

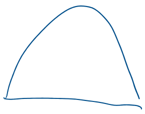
Tcheby says that $\geq 89\%$ of measurements lie between 20 and 80

k	$1 - 1/k^2$
1.5	5/9
2	3/4
2.5	21/25
3	8/9

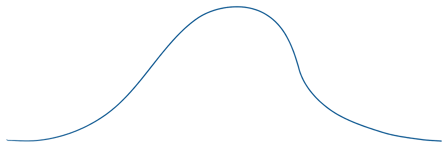
so $\geq 56\%$ of points lie within 1.5 std dev of mean
$\geq 75\%$ " 2 "
$\geq 84\%$ 2.5
$\geq 89\%$ 3

The Empirical Rule: only works for data sets that are unimodal and

that are unimodal and
roughly symmetrical



"mound"



"bell"

approximately 68% of measurements lie with 1 std dev of mean
" 95% " 2
99.7% 3