Section 2.3: Tchebysheff's Theorem and the Empirical Rule

Tuesday, October 29, 2019 10:54 AM

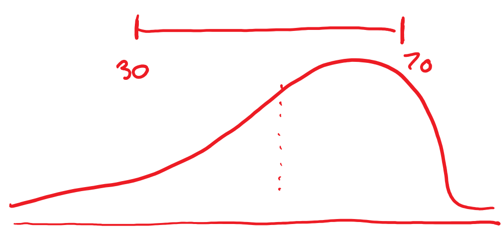
or Chebyshev

Tchebysheff's theorem: works for all distributions/ data sets

(symmetrical or skewed, unimodal or multimodal)

- for any set of measurements, at least $\left(1 - \frac{1}{k^2}\right)$ of the measurements fall within k standard deviations of the mean for $k \geq 1$

example:



$\bar{x} = 50$

$s = 10$

look at the interval from 30 to 70

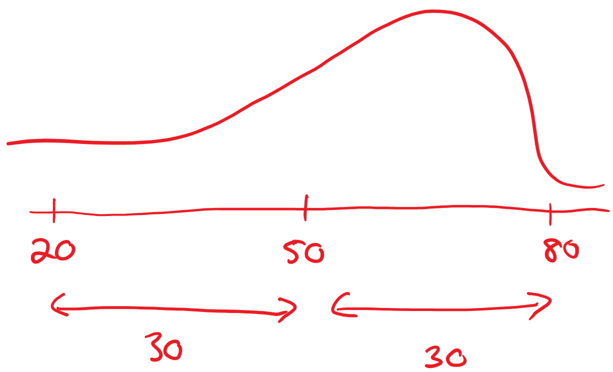
30 is two standard deviations below the mean, and 70 is two standard deviations above

so $k = 2$

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

Tcheby says at least 75% of the measurements fall between 30 and 70

what about the number of measurements between 20 and 80?



= 3 std dev

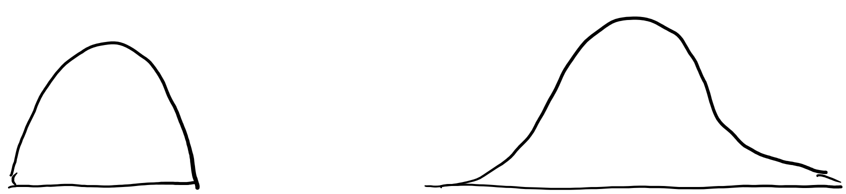
so $$1 - \frac{1}{k^2} = 1 - \frac{1}{3^2} = \frac{8}{9} \text{ or } 88.\bar{8}\%$$

and at least 89% of the measurements fall within the interval 20 to 80

k	$1 - 1/k^2$			
1	0	so $\geq 0\%$ *	lie within	$\bar{x} \pm 1s$
1.5	$5/9$	$\geq 55.\bar{5}\%$		$\bar{x} \pm 1.5s$
2	$3/4$	$\geq 75\%$		$\bar{x} \pm 2s$
2.5	$21/25$	$\geq 84\%$		$\bar{x} \pm 2.5s$
3	$8/9$	$\geq 88.\bar{8}\%$		$\bar{x} \pm 3s$

* $\geq 0\%$ may seem like a totally useless statement, but what it says is that it is possible to have no data within one std dev of the mean

The Empirical Rule: only works for "mound-shaped" or "bell-shaped" curves



mound   bell

– for unimodal and <u>roughly</u> symmetrical

approximately 68% of the measurements fall within $\bar{x} \pm 1s$

" 95% $\bar{x} \pm 2s$

" 99.7% $\bar{x} \pm 3s$