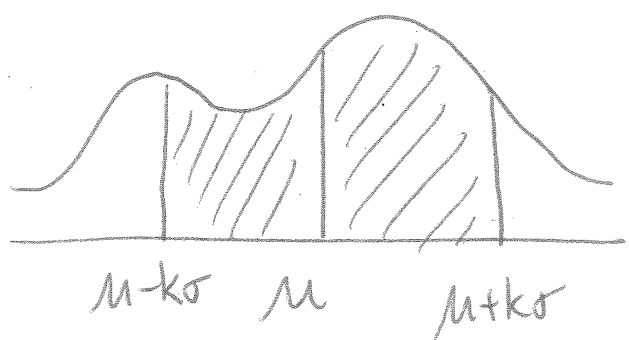
6.3 Tchebysheff and Empirical Rules

Tchebysheff's Rule

For any data set:
At least $(1-\frac{1}{k^2})$ of the measurements lie within the interval $\mu-k\sigma \le x \le \mu+k\sigma$, where k is any real number >1.



Ex: The mean of a data set is 25 and the standard deviation is 4. What percentage of measurements will fall between 14.2 and 35.8?

$$\mu+k\sigma = 35.8$$
$$25+k(4) = 35.8$$
$$4k = 10.8$$
$$k = 2.7$$

$$1-\frac{1}{(2.7)^2} \approx 0.86$$

At least 86%

<u>Ex</u>: The average age of programmers at a Company is 31, with a standard deviation of 5. Find the range in which at least 56% of the ages fall.

$$1 - \frac{1}{k^2} = 0.56$$

$$0.44 = \frac{1}{k^2}$$

$$0.44k^2 = 1$$

$$k^2 = \frac{1}{0.44}$$

$$k = \sqrt{\left(\frac{1}{0.44}\right)}$$

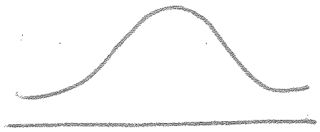
$$k \approx 1.5$$

$$\mu - k\sigma \leq x \leq \mu + k\sigma$$

$$31 - 1.5(5) \leq x \leq 31 + 1.5(5)$$

$$23.5 \leq x \leq 38.5$$

A mound-shaped (or bell-shaped) data set is unimodal and symmetric:



The Empirical Rule

If the data is approximately mound-shaped then:

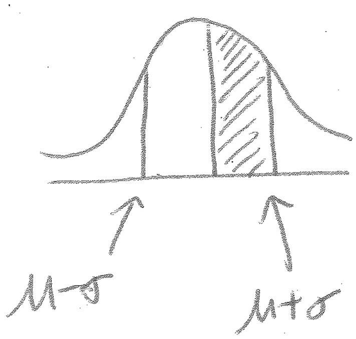
$\mu - \sigma \le x \le \mu + \sigma$ contains approximately 68% of data

$\mu - 2\sigma \le x \le \mu + 2\sigma$ " 95% "

$\mu - 3\sigma \le x \le \mu + 3\sigma$ " 99.7% "

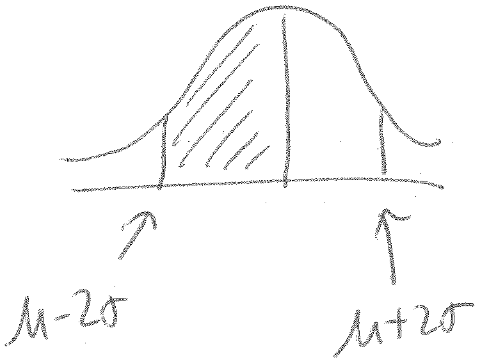
Ex: Given a mound-shaped data set, estimate the % of data in the shaded region.

a)



$\frac{0.68}{2} = 0.34$ or 34%

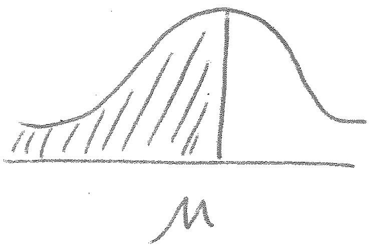
b)



$\frac{0.95}{2} = 0.475$

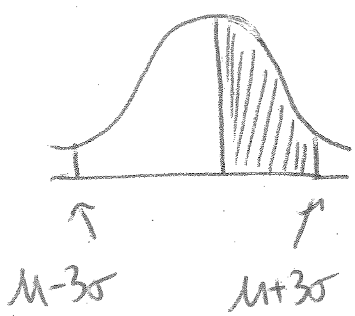
or 47.5%

c)



0.5 or 50%

d)

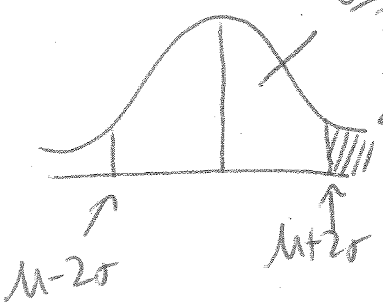


$\frac{0.997}{2} = 0.4985$ or 49.85%

e)

$\frac{0.95}{2} = 0.475$

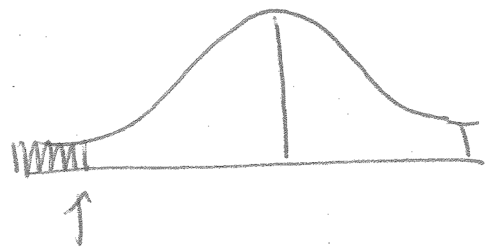
$0.5 - 0.475$



0.025 or 2.5%

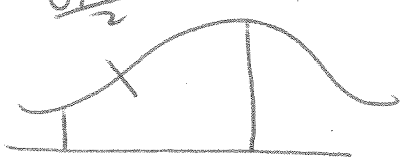
Ex: At a software company, hours worked last week are mound-shaped with a mean of 42 and a standard deviation of 2. What % of employees worked less than 36 hours (approximately)?

$$36 = \mu - k\sigma$$
$$36 = 42 - k(2)$$
$$-6 = -2k$$
$$3 = k$$

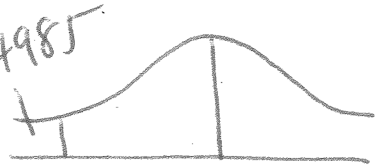


$\mu - 3\sigma$

$\frac{0.997}{2} = 0.4985$



$0.5 - 0.4985$



0.0015 or 0.15%