

Section 3.4: cont'd

Tuesday, May 07, 2013
11:34 AM

Assignment #2 due on
Tuesday, May 21st

Quiz #2 on
Friday, May 24th

on Sections 31.9 + 31.10 of DE
and Chapters 1 to 3 of Stats

Formula sheet posted later this week

Recall: z-score

$$z = \frac{x - \mu}{\sigma} \quad \text{for population}$$

$$= \frac{x - \bar{x}}{s} \quad \text{for samples}$$

What does Tchebysheff say?

at least 75% of all measurements
will have a z-score between -2 and 2

Example: 0.11 (11% of cases)

Empirical Rule (mound-shaped):

~ 95% of all measurements will have z-scores between -2 and 2

so the point is:

$$|z| > 2$$

- for any distribution, will happen $< 25\%$ of the time
- for mound-shaped, will happen $\sim 5\%$ of the time

so $|z| > 2$ is somewhat unlikely

$$|z| > 3$$

- for any distribution, will happen $< 10\%$ of the time
- for mound-shaped, will happen $\sim 0.3\%$ of the time

so $|z| > 3$ is very unlikely

so outliers with $|z| > 3$ need to be examined carefully

- may be an error in measurement
 - but may still be valid
-

example:

Ten DVDs were picked at random from Pat's DVD collection, and the number of Oscars won by each movie was tabulated:

0, 0, 1, 2, 0, 0, 2, 0, 11, 1

mean $\bar{x} = 1.70$

std dev $s = 3.37$

Calculate the z-score for any outliers and state your conclusions.

$$Z = \frac{x - \bar{x}}{s} = \frac{11 - 1.70}{3.37} = 2.761158$$
$$= 2.76$$

this z-score is unlikely

(note: LOTR: Return of the King was not a typical movie)