Section 2.4: Measures of Relative Standing

Wednesday, October 30, 2019 11:03 AM

measures of relative standing

- give an indication of the position of an individual data point with respect to the rest of the data

example: you ran a race and your time was 1 hour and 56 minutes

but how did everyone else do?

measures: percentiles (we will not cover)

z-score

---

z-score:

for populations, it's

$$z = \frac{x - \mu}{\sigma}$$

where x = data point of interest
μ = population mean
σ = population standard deviation

for samples, it's

$$z = \frac{x - \bar{x}}{s}$$

where $\bar{x}$ = sample mean
s = sample standard deviation

z is then the number of standard deviations above the mean that the data point is

- if z is negative, then the data point is below the mean

2019/10/31

so, how likely are various z-scores?

notation: $|z|$ is the absolute value of z (how far from the origin z is)

$|3| = 3$
$|-3| = 3$
$|0| = 0$

$|z| > 2$ for any distribution, will happen $\leq 25\%$ of the time (Tcheby)

for mound-shaped, will happen
~ 5% of the time
(~2.5% for $z > 2$,
~2.5% for $z < -2$)

so $|z| > 2$ is somewhat unlikely

$|z| > 3$ - for any distribution, will happen
$\leq 11\%$ of the time (Tcheby)

- for mound-shaped, will happen
~0.3% of the time
(~0.15% above, ~0.15% below)

so $|z| > 3$ is very unlikely

example: Ten DVDs were picked at random from Pat's collection, and the number of Oscars (awards) won by each movie was recorded.

results: 0, 0, 1, 2, 0, 0, 2, 0, 11, 1

The standard deviation for this data set is 3.37. Calculate the z-score for any outliers and state whether those data points are likely or unlikely.

answer: $\bar{x} = 1.7$
$s = 3.37$

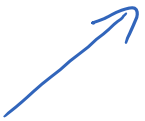
$$z = \frac{x - \bar{x}}{s}$$

$$= \frac{11 - 1.7}{3.37}$$

$$= 2.761158$$

$$= 2.76$$

(in general, we round z-scores to 2 decimal places)



unlikely

note: Lord of the Rings: Return of the King was not a typical movie