



DataCamp

Learning by doing

Variability

- A measure that describes the range and diversity of scores in a distribution
 - *Standard deviation* (SD): the average deviation from the mean in a distribution
 - *Variance* = SD^2

Variability

- *Variance = SD²*

$$SD^2 = [\Sigma(X - M)^2] / N$$

Variance

- Variation is natural and observed in all species and that's good:
 - *On the Origin of Species* (1859)
 - *Variation Under Domestication* (1868)

Linsanity!



Jeremy Lin (10 games)

Points per game	(X-M)	(X-M) ²
28	5.3	28.09
26	3.3	10.89
10	-12.7	161.29
27	4.3	18.49
20	-2.7	7.29
38	15.3	234.09
23	0.3	0.09
28	5.3	28.09
25	2.3	5.29
2	-20.7	428.49
M = 227/10 = 22.7	M = 0/10 = 0	M = 922.1/10 = 92.21

Results

- $M = \text{Mean} = 22.7$
- $SD = \text{Standard Deviation} = 9.6$
- $SD^2 = \text{Variance} = 92.21$

Notation

- M = Mean
- SD = Standard Deviation
- SD^2 = Variance (also known as MS)
 - MS stands for Mean Squares
 - SS stands for Sum of Squares

Summary statistics: Review

- Important concepts
 - Central tendency (mean, median, mode)
 - Variability (standard deviation and variance)

Summary statistics: Review

- Summary statistics (formulae to know)
 - $M = (\Sigma X) / N$
 - $SD^2 = [\Sigma(X - M)^2] / N$
 - Used for descriptive statistics
 - $SD^2 = [\Sigma(X - M)^2] / (N - 1)$
 - Used for inferential statistics