STAT 231

Roadmap

- Measures of Dispersion
- ~~Method~~ MEASURE OF SKEWNESS
- KURTOSIS
- Measures of Association.

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \text{S. M}$$

$$s^2 = \frac{1}{n-1}\left[\sum (y_i - \bar{y})^2\right]$$

= SAMPLE VARIANCE

s = Positive square root of s^2

Properties

Another equivalent way

$$s^2 = \frac{1}{n-1}\left[\sum y_i^2 - n\bar{y}^2\right]$$

Example: $\{y_1, \ldots y_{20}\}$

Mean $= 25$. $s^2 = 625$

One observation $y_{20} = 44$ is wrong and we want to discard it.

Find the new mean and the new s.d.

New mean ?

Old Total: $n \times \bar{y}_{old} = 20 \times 25$

$= 500.$

New Total: $500 - 44 = 456$

New mean $\bar{y}_{new} = 456/19 = 24$

New variance ?

OLD DATA SET

$$625 = \frac{1}{19}\left[\sum_{i=1}^{20} y_i^2 - 20 \times 25^2\right]$$

$$\sum_{i=1}^{20} y_i^2 = [625 \times 19 + 20 \times 25^2]$$

X

New sum of squares = $X - 44^2$.

New $s^2 = \frac{1}{n-1}\left[\sum_{i=1}^{19} y_i^2 - n\,\bar{y}_{new}^2\right]$

(n → 19)

If the data set is normally distributed,

approximately 68% lie between

$(\bar{y} - s,\ \bar{y} + s)$

95% lie within

$(\bar{y} - 2s,\ \bar{y} + 2s)$

Clicker question 1

- $\{y_1, \ldots y_n\}$ incomes
  $\$\varepsilon$ is transferred by Robin Hood from the richest to the poorest.

- What happens to the mean ?

  (1) • increases

  (2) • decreases

  (3) • stays the same ? ✓

Clicker Q 2

- What happens to the variance?

(1) • increases

(2) • decreases

(3) • stays the same

(4) not enough info .

$$\{y_1, \ldots y_n\}$$
$$\downarrow$$
$$\bar{y}$$

Mean does not change.

$$n\bar{y} - \bar{y} = \text{New Total}$$

$$\frac{(n-1)\bar{y}}{n-1} =$$

$$\bar{y}$$

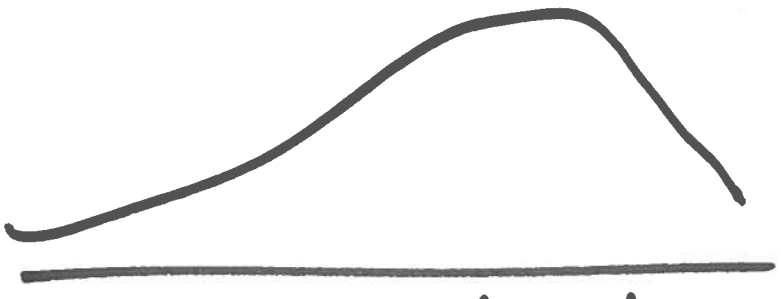
$$\frac{1}{n-1}\left[\sum(y_i-\bar{y})^2\right]$$

Measures of Symmetry

symmetric

right skewed

+vely skewed.



negatively skewed.

left skewed.

1, 3, 5, 7, 9 10^7

$\underset{\times}{}$

Mean = Median.

Asymmetry = Mean − Median.

$< 0 \Rightarrow$ evidence of a left skew

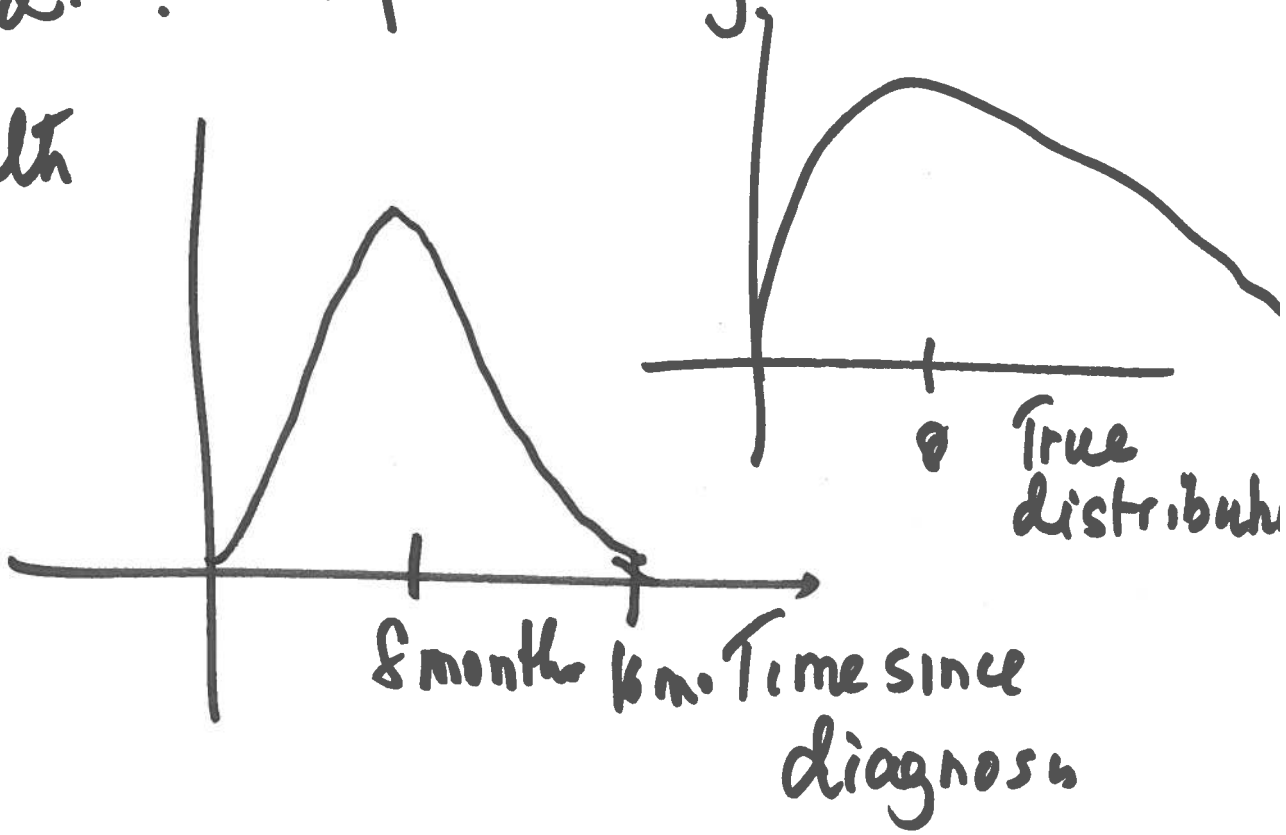
$> 0 \Rightarrow$ evidence of a right-skewed distribution

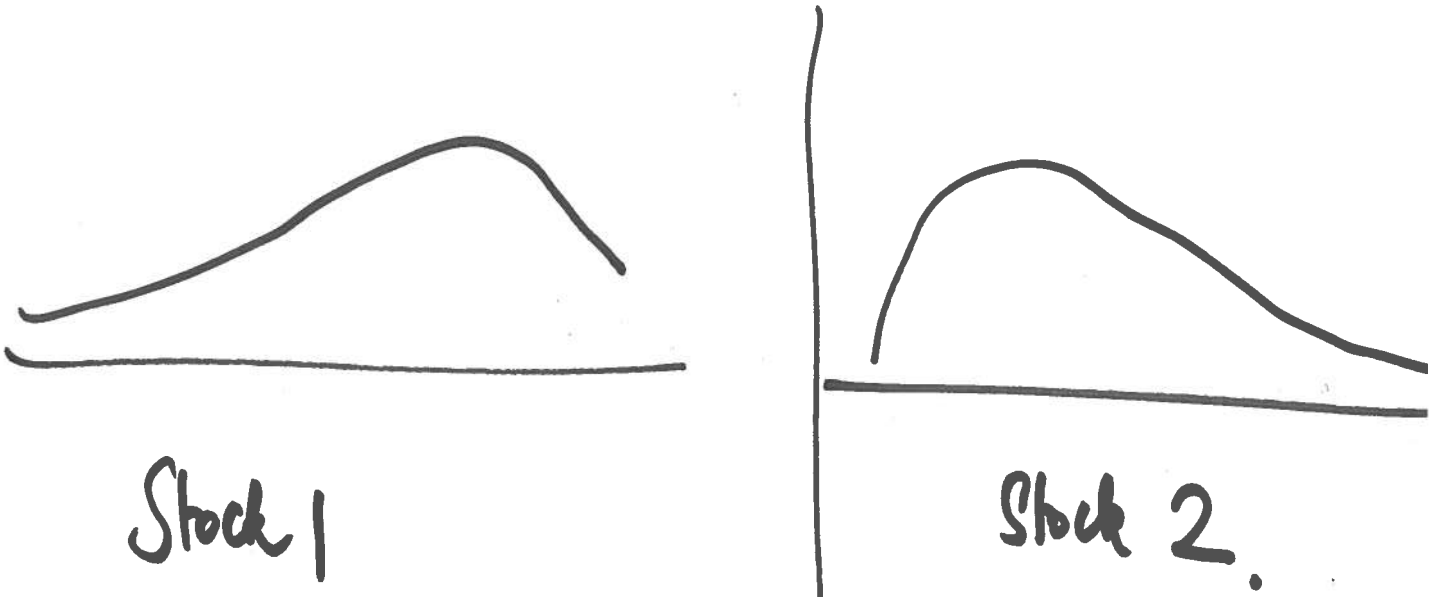
$= 0 \Rightarrow$ evidence of symmetry

$$\text{Skewness} = \frac{\frac{1}{n}\sum(y_i - \bar{y})^3}{\left[\frac{1}{n}\sum(y_i - \bar{y})^2\right]^{3/2}}$$

Gould. : Stepen Jay,

Death



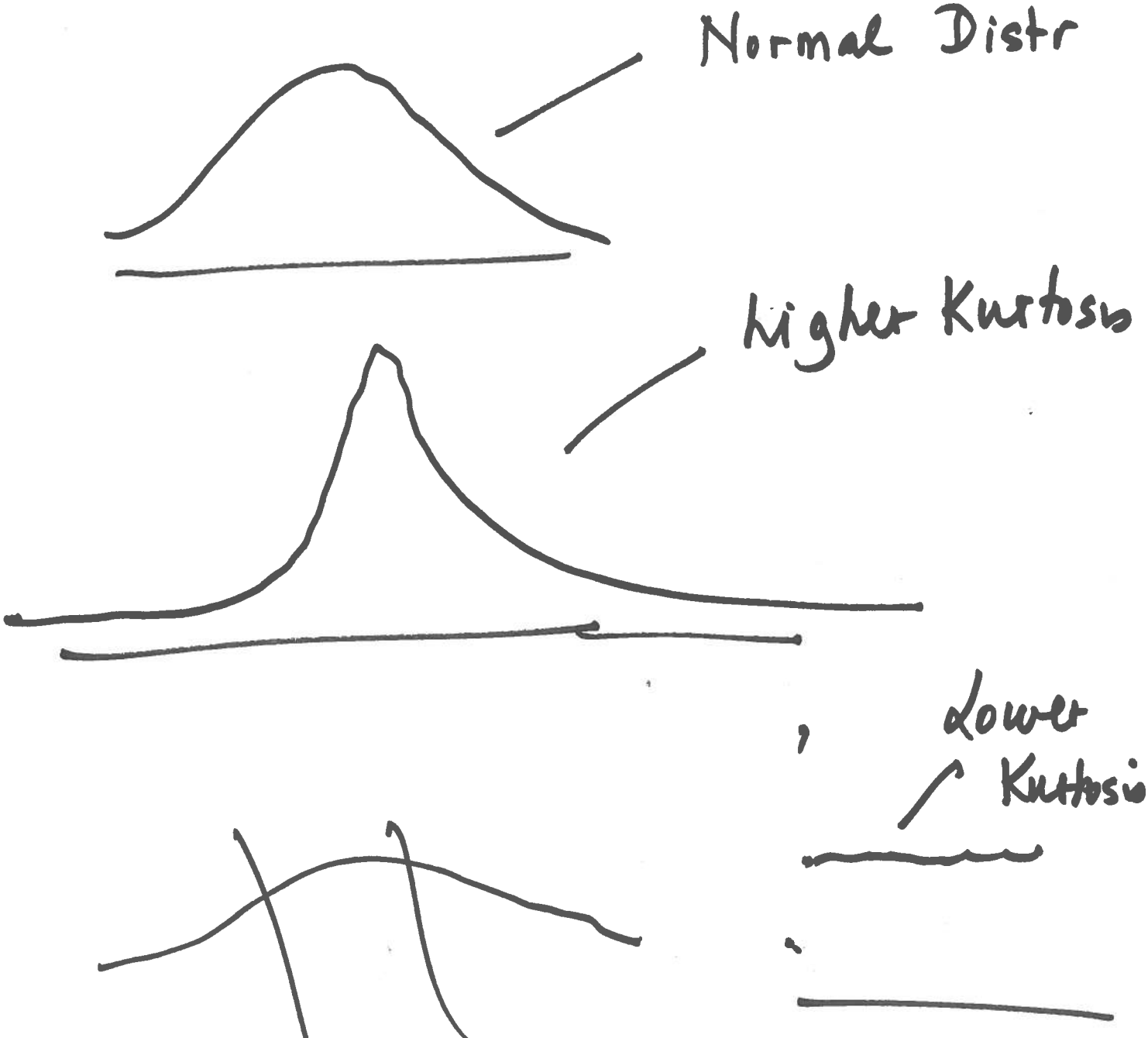


Stock 1 Stock 2.

These stocks, even though they have the same variance, are not equivalent.

KURTOSIS

Compares the frequency of extreme observations & with the Normal Distribution (how fat the tails are)


Normal Distr
Higher Kurtosis
Lower Kurtosis

For a Normal Distribution,

$$K = 3$$

$K >> 3 \Rightarrow$ more extreme obs. than normal

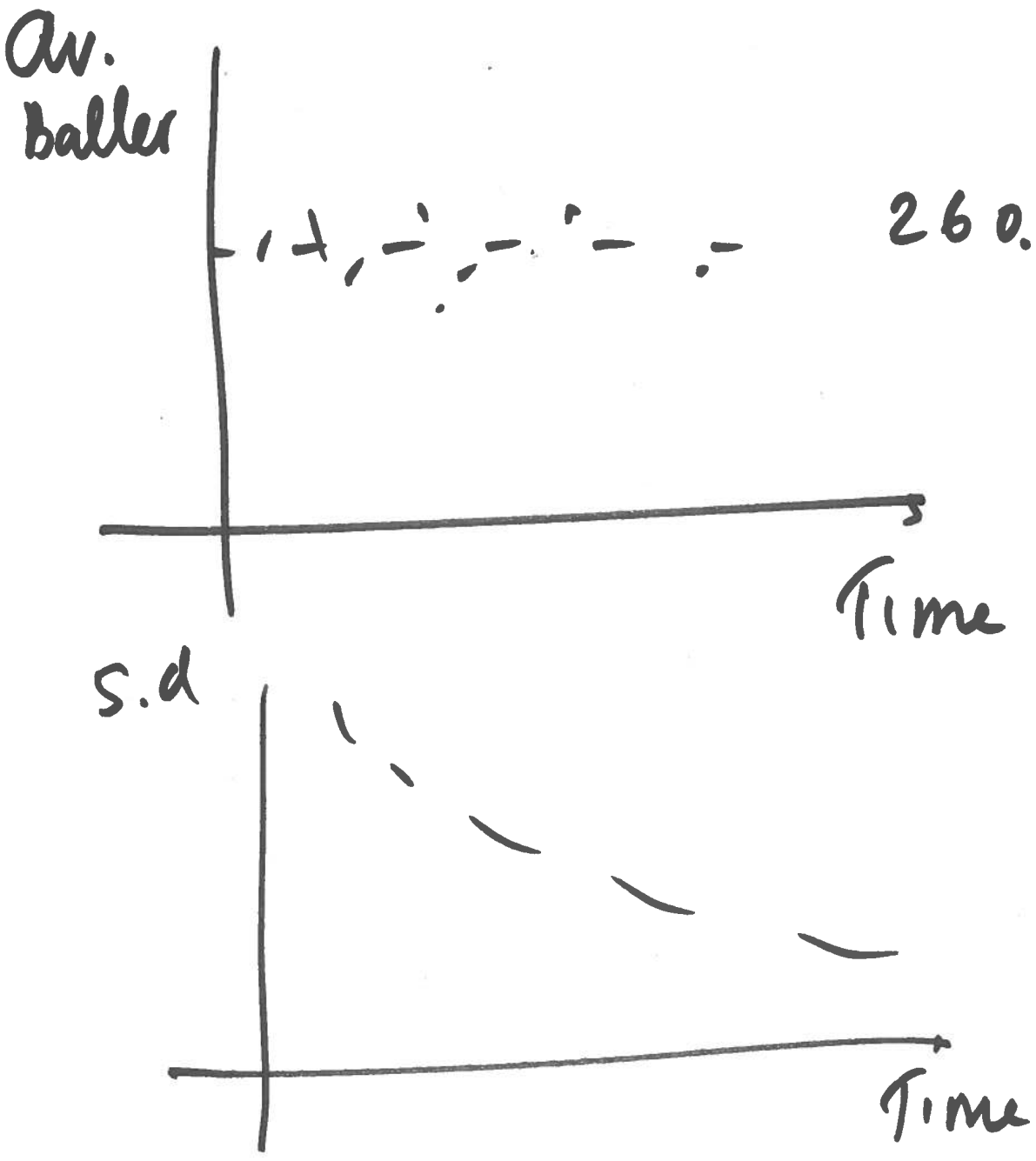
$K << 3 \Rightarrow$ Less extreme obs. than the normal.

Gould. → "FULL HOUSE"

Why the 400 hitter disappeared?

- Absolute
- Relative

} Batters have gotten worse.


Av.
Baller
260.
Time
S.d
Time