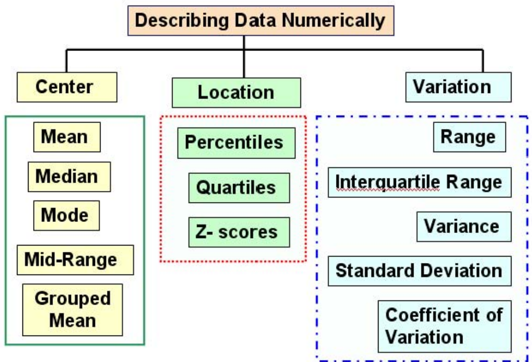
DUXBURY
THOMSON LEARNING
Professor David C
Rochester Institute of
Dept of Mathematics

Section 4.4 – Z-Scores and the Empirical Rule



DUXBURY
THOMSON LEARNING
Professor David Crystal
Rochester Institute of Technology
Dept of Mathematics & Statistics

GPA Example

A sample of GPAs of 40 freshman college students appear below (sorted in increasing order)

1.40	1.90	1.90	2.00	2.10	2.10	2.20	2.30
2.30	2.40	2.50	2.60	2.60	2.70	2.80	2.80
2.90	2.90	2.90	2.90	3.00	3.00	3.00	3.00
3.10	3.10	3.20	3.20	3.30	3.30	3.40	3.40
3.50	3.50	3.60	3.70	3.80	3.80	3.90	4.00

- Enter data in TICalc in L1
- Set Window →

  Xmin= 1 Ymin=-1
  Xmax= 5 Ymax=20
  Xscl= .5 Yscl= 5
- Check Stats **Xbar=2.9 s=.62017**
- Create a Box Plot and Histogram to review the distribution of the data

DUXBURY

THOMSON LEARNING

Professor David Crystal
Rochester Institute of Technology
Dept of Mathematics & Statistics

Shape of the RAW GPA Data

- Sketch the GPA Box Plot and Histogram and review the distribution of the data.

- Is the distribution of the data reasonably symmetric and unimodal?

3

DUXBURY

THOMSON LEARNING

Professor David Crystal
Rochester Institute of Technology
Dept of Mathematics & Statistics

Z-Scores and Location

- By itself, a raw data value provides very little information about how that particular score compares with other values in the distribution.
- If the raw score is transformed into a z-score, however, the value of the z-score tells exactly where the score is located relative to all the other scores in the distribution.

The formula used with sample data is

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

- Enter Z-Scores into L2
- Use TICalc (Vars) option 5:Statistics when entering formula into L2
- Set Window →
- Check Stats Xbar=0 s=1
- Create a Box Plot and Histogram to review the distribution of the data

Xmin= -3 Ymin=-1
Xmax= 3 Ymax=30
Xscl= 1 Yscl= 20

4

DUXBURY
THOMSON LEARNING

Professor David Crystal
Rochester Institute of Technology
Dept of Mathematics & Statistics

Z-Scores

Computing the z score is often referred to as **standardization** and the z score is called a **standardized score**.

The formula used with sample data is

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

- Enter Z-Scores into L2
  - Use TICalc (Vars) - 5:Statistics when entering formula into L2.

DUXBURY
THOMSON LEARNING
Professor David Crystal
Rochester Institute of Technology
Dept of Mathematics & Statistics

Shape of the GPA ZScores

- Sketch the GPA Box Plot and Histogram and review the distribution of the data.

- Compare the RAW and ZScores, how do their distributions differ?

Copyright © 2005 Brooks/Cole, a division of Thomson Learning, Inc.

DUXBURY
THOMSON LEARNING
Professor David Crystal
Rochester Institute of Technology
Dept of Mathematics & Statistics

Z-Scores (continued)

The **z score** corresponding to a particular observation in a data set is

$$\text{zscore} = \frac{\text{observation} - \text{mean}}{\text{standard deviation}}$$

- Now Count the Z-Scores and find the % in each group:
  - Zscores +/- 1 std dev:______/40= ______%
  - Zscores +/- 2 std dev :______/40= ______%
  - Zscores +/- 3 std dev :______/40= ______%

- *The z score is how many standard deviations the observation is from the mean.*
- *A positive z score indicates the observation is above the mean and a negative z score indicates the observation is below the mean.*

DUXBURY
THOMSON LEARNING
Professor David Crystal
Rochester Institute of Technology
Dept of Mathematics & Statistics

Z-Scores by Number of Standard Deviations

Interval	# GPA Z-scores	Empirical Rule
within 1 standard deviation of the mean	27/40 = 67.5%	≈ 68%
within 2 standard deviations of the mean	39/40 = 97.5%	≈ 95%
within 3 standard deviations of the mean	40/40 = 100%	≈99.7%

Now let's look at the empirical rule and decide if our estimates are reasonable? →

Copyright © 2005 Brooks/Cole, a division of Thomson Learning, Inc.

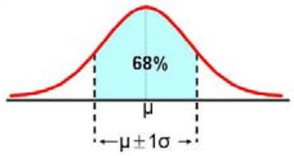
DUXBURY
THOMSON LEARNING
Professor David Crystal
Rochester Institute of Technology
Dept of Mathematics & Statistics

Several Rules related to standard deviation

"The Empirical Rule"

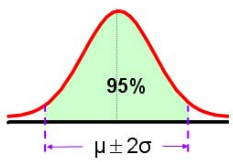
1. Also called (in the text) the "68 – 95 – 99.7 Rule"

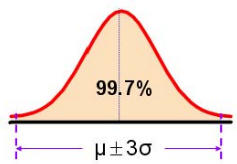
$\mu \pm 1\sigma$ contains about 68% of the values in the population or the sample



$\mu \pm 2\sigma$ contains about 95% of the values in the population or the sample

$\mu \pm 3\sigma$ contains about 99.7% of the values in the population or the sample

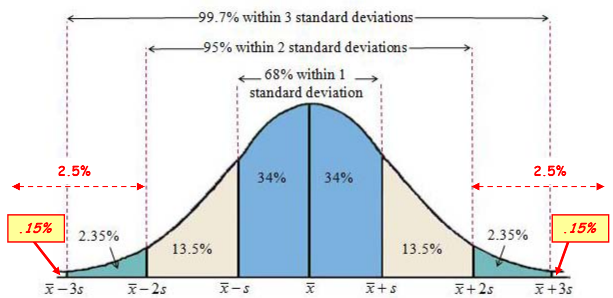




DUXBURY
THOMSON LEARNING
Professor David Crystal
Rochester Institute of Technology
Dept of Mathematics & Statistics

The Empirical Rule "68-95-99.7"

Next we put all three statements in a single figure below.



2. Note this "rule" is based on the bell-shaped "normal distribution", but often applies approximately for other shapes.

DUXBURY
THOMSON LEARNING

Professor David Crystal
Rochester Institute of Technology
Dept of Mathematics & Statistics

Z-Scores and Location **Summary**

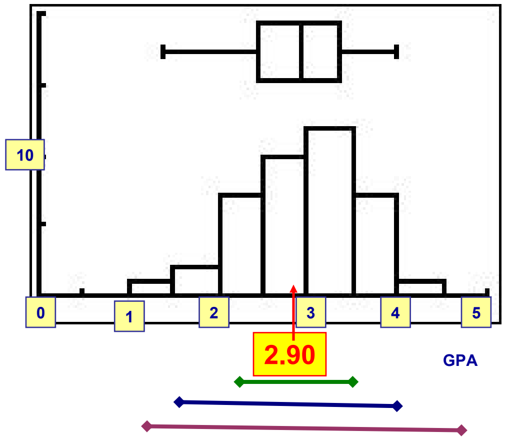
- The process to calculate Z scores is called standardization and the Z-Score is called a standardized score. *It is a position marker.*
- The process of changing a raw data value into a z-score involves creating a signed number, such that
  a) The sign of the z-score (+ or –) identifies whether the data value is located above the mean (positive) or below the mean (negative).
  b) The numerical value of the z-score corresponds to the number of standard deviations between the data value and the mean of the distribution.
- **Also, the terms in the formula can be regrouped to create an equation for computing the data value corresponding to any specific z-score.**

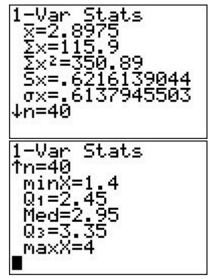
$$X = \mu + z\sigma$$

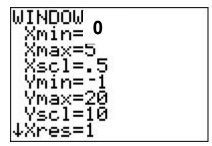
Appendix A

	GPA	Zscore		GPA	Zscore		GPA	Zscore		GPA	Zscore
1	1.4	-2.4	8	2.3	-1.0	22	3.0	0.2	35	3.6	1.1
2	1.9	-1.6	9	2.3	-1.0	23	3.0	0.2	36	3.7	1.3
3	1.9	-1.6	10	2.4	-0.8	24	3.0	0.2	37	3.8	1.5
4	2.0	-1.4	11	2.5	-0.6	25	3.1	0.3	38	3.8	1.5
5	2.1	-1.3	12	2.5	-0.6	26	3.1	0.3	39	3.9	1.6
6	2.1	-1.3	13	2.6	-0.5	27	3.2	0.5	40	4.0	1.8
7	2.2	-1.1	14	2.7	-0.3	28	3.2	0.5			
			15	2.8	-0.2	29	3.3	0.6	Mean	2.90	0.00
			16	2.8	-0.2	30	3.3	0.6	std dev	0.62	1.00
			17	2.9	0.0	31	3.4	0.8			
			18	2.9	0.0	32	3.4	0.8			
			19	2.9	0.0	33	3.5	1.0			
			20	2.9	0.0	34	3.5	1.0			
			21	3.0	0.2						

Appendix B: "Raw Data"



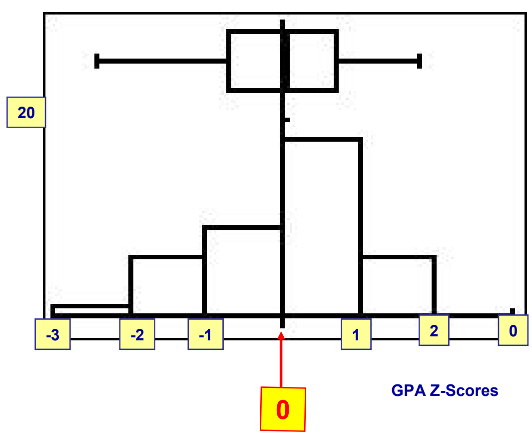


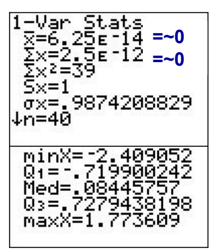


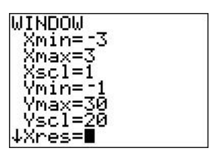
2.90 ± 1(.62) → 2.28 to 3.52
(27/40=67.5%)
2.90 ± 2(.62) → 1.66 to 4.14 (39/40=97.5%)
2.90 ± 3(.62) → 1.04 to 4.76 (40/40=100%)

13

Appendix C "Standardized Z-Scores"





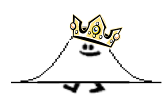


14

4.4B Notes

Understanding Z-Score and the Empirical Rule (68-95-99.7)

What's my area?



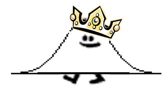
Input the following command into a graphing calculator in order to graph a normal curve with a mean of 20 and standard deviation of 3.

Y1 = normalpdf(X,20,3)   (Window x: [10,30] y: [0,0.2])

Use the command 2nd trace, 7 to find the area under the curve for the: (Round to 3 decimal places.)

1) Lower limit: 17   Upper limit: 23 Area: _________
2) Lower limit: 14   Upper limit: 26 Area: _________
3) Lower limit: 11   Upper limit: 29 Area: _________

What's my area?



Graph a normal curve with a mean of 50 and standard deviation of 5.

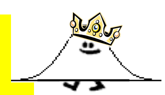
Y1 = normalpdf(X,50,5)   (x: [30,70] y: [0,0.1])

Find the area under the curve for the following:

1) Lower limit: 45   Upper limit: 55 Area: ________
2) Lower limit: 40   Upper limit: 60 Area: ________
3) Lower limit: 35   Upper limit: 65 Area: ________

What pattern do you notice?

What's my area? Pattern you should have noticed was the Empirical Rule.

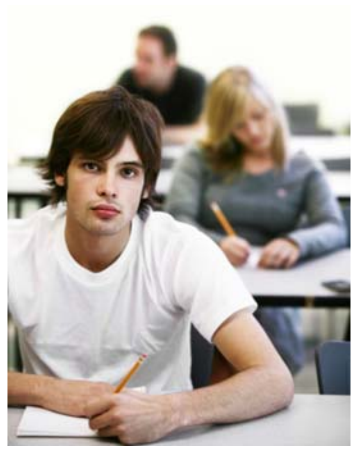


Approximately 68% of the observations are within 1 standard deviation of the mean

Approximately 95% of the observations are within 2 standard deviation of the mean

Approximately 99.7% of the observations are within 3 standard deviation of the mean

The height of male students at PWSH is approximately normally distributed with a mean of 71 inches and standard deviation of 2.5 inches.



a) What percent of the male students are shorter than 66 inches?

About 2.5%

b) Taller than 73.5 inches?

About 16%

c) Between 66 & 73.5 inches?

About 81.5%

19

Measures of Relative Standing

Z-score

A z-score tells us how many standard deviations the value is from the mean.

$$\text{z - score} = \frac{\text{value - mean}}{\text{standard deviation}}$$

One example of *standardized* score.

20

What do these *z*-scores mean?

-2.3 2.3 standard deviations below the mean

1.8 1.8 standard deviations above the mean

-4.3 4.3 standard deviations below the mean

Sally is taking two different math achievement tests with different means and standard deviations. The mean score on test A was 56 with a standard deviation of 3.5, while the mean score on test B was 65 with a standard deviation of 2.8. Sally scored a 62 on test A and a 69 on test B. On which test did Sally score the best?

Z-score on test A

$$z = \frac{62 - 56}{3.5} = 1.714$$

Z-score on test B

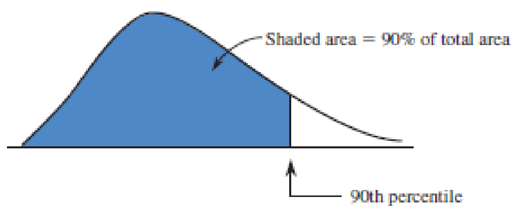
$$z = \frac{69 - 65}{2.8} = 1.429$$

She did better on test A.

Measures of Relative Standing

Percentiles

A **percentile** is a value in the data set where *r* percent of the observations fall **AT or BELOW** that value



23

In addition to weight and length, head circumference is another measure of health in newborn babies. The National Center for Health Statistics reports the following summary values for head circumference (in cm) at birth for boys.

Head circumference (cm)	32.2	33.2	34.5	35.8	37.0	38.2	38.6
Percentile	5	10	25	50	75	90	95

What percent of newborn boys had head circumferences greater than 37.0 cm? 25%

10% of newborn babies have head circumferences bigger than what value? 38.2 cm

24