4N Introduction to the Normal Distribution

Describing Location in a Distribution and
and Density Curves

A) More Practice With Standardized Scores

Review Z-Scores

- One way to describe relative position in a data set is to tell how many standard deviations above or below the mean the observation is.

Standardized Value: "z-score"
If the mean and standard deviation of a distribution are known, the "z-score" of a particular observation, x, is:

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

- ***SAMPLE DATA: Consider the following test scores for a small class:***

79	81	80	77	73	83	74	93	78	80	75	67	73
77	83	86	90	79	85	83	89	84	82	77	72	

1) What is the distribution of the test scores?_____

```
6 |  7
7 | 2334
7 | 5777899
8 | 00123334
8 | 569
9 | 03
```

Minitab

Session

Descriptive Statistics: Test 1 scores

Variable	N	Mean	Median	TrMean	StDev	SE Mean
Test 1 scores	25	80.00	80.00	80.00	6.07	1.21

Variable	Minimum	Maximum	Q1	Q3
Test 1 scores	67.00	93.00	76.00	83.50

2) Reading the Minitab output, the mean is ___ and standard deviations is ____.

Answer the following questions and use Z-Scores to justify your answers:

1. Julia's score was 86. How did she perform on this test relative to her peers? Her score is "above average"...but how far above average is it?

2. Kevin's score was 72. How did he perform on this test relative to her peers?

3. Katie's score was 80. How did she perform on this test relative to her peers?

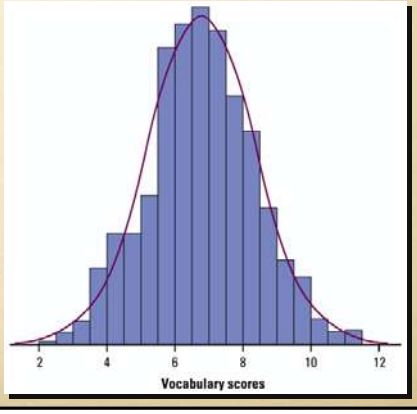
Comparing Scores – More Practice Problems

- Standardized values can be used to compare scores from two different distributions.
  - ☑ Statistics Test: mean = 80, std dev = 6.07
  - ☑ Chemistry Test: mean = 76, std dev = 4
  - ☑ Jenny got an 86 in Statistics and 82 in Chemistry.
  - ☑ On which test did she perform better?

B) Density Curve

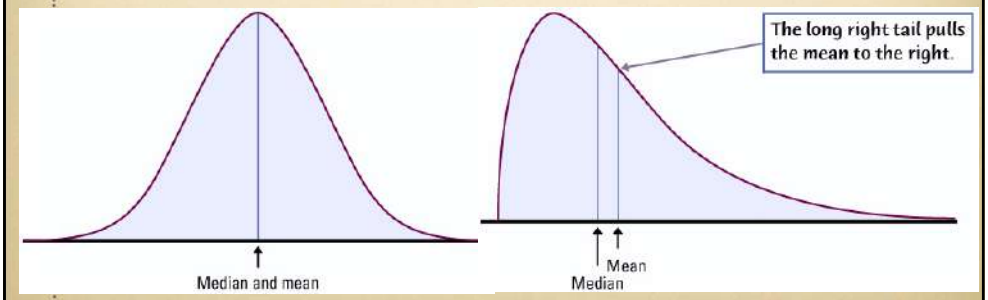
- In Chapter 1, you learned how to plot a dataset to describe its shape, center, spread, etc.
  - Sometimes, the overall pattern of a large number of observations is so regular that we can describe it using a smooth curve.

Density Curve:
An idealized description of the overall pattern of a distribution.
Area underneath = 1, representing 100% of observations.



Density Curves

- Density Curves come in many different shapes; symmetric, skewed, uniform, etc.
- The area of a region of a density curve represents the % of observations that fall in that region.
- The median of a density curve cuts the area in half.
- The mean of a density curve is its "balance point."



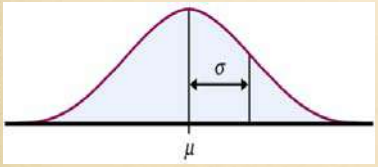
Summary

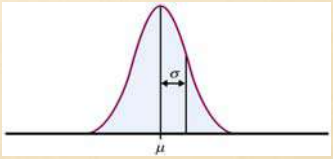
- We can describe the overall pattern of a distribution using a density curve.
- The area under any density curve = 1. This represents 100% of observations.
- Areas on a density curve represent % of observations over certain regions.
- An individual observation's relative standing can be described using a z-score.

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

B) Normal Distributions

- Normal Curves: symmetric, single-peaked, bell-shaped. μ and median are the same. Size of the σ will affect the spread of the normal curve.





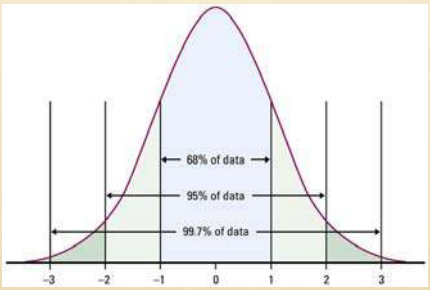
Assessing Normality

- **For now,** construct a dot plot, stem plot, histogram, see if graph is approximately bell-shaped and symmetric. Median and Mean should be close. Then mark off the -2, -1, +1, +2 SD points and check the 68-95-99.7 rule.

The 68–95–99.7 Rule

In the Normal distribution with mean μ and standard deviation σ:

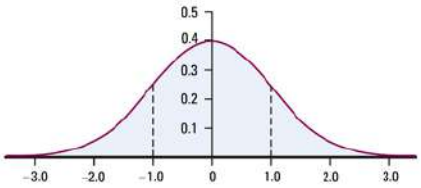
- Approximately 68% of the observations fall within σ of the mean μ.
- Approximately 95% of the observations fall within 2σ of μ.
- Approximately 99.7% of the observations fall within 3σ of μ.



Standard Normal Distribution

The **standard Normal distribution** is the Normal distribution N(0, 1) with mean 0 and standard deviation 1 (Figure 2.15).

Figure 2.15 *Standard Normal distribution.*



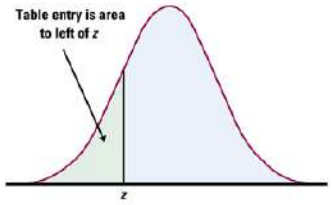
If a variable x has any Normal distribution N(μ, σ) with mean μ and standard deviation σ, then the standardized variable

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

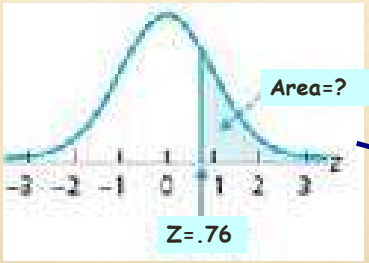
has the standard Normal distribution.

The Standard Normal Table

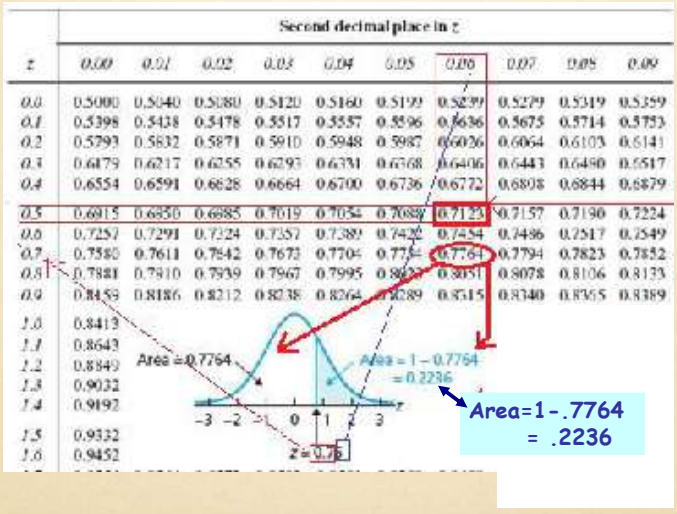
Table A is a table of areas under the standard Normal curve. The table entry for each value z is the area under the curve to the left of z.



Example A



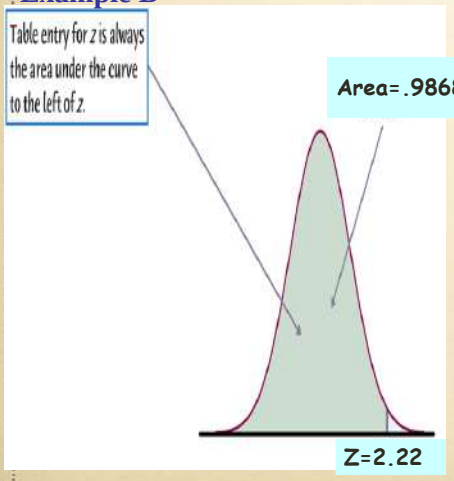
See Green AP Exam Table A



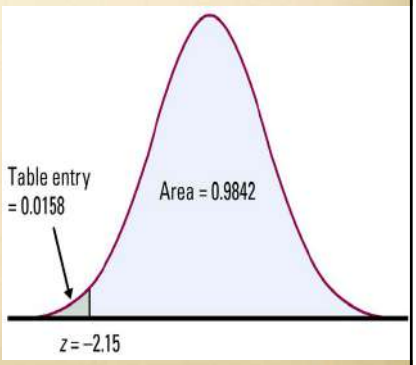
Use Green Tables to find the Area under the curve

Are you getting the correct Areas?

Example B



Example C



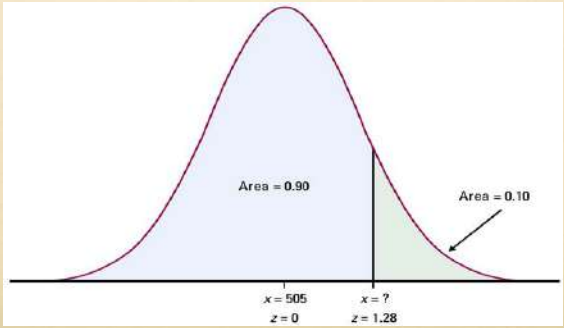
Solving Problems Involving Normal Distributions

Step 1: *State the problem* in terms of the observed variable x. Draw a picture of the distribution and shade the area of interest under the curve.

Step 2: *Standardize and draw a picture.* Standardize x to restate the problem in terms of a standard Normal variable z. Draw a picture to show the area of interest under the standard Normal curve.

Step 3: *Use the table.* Find the required area under the standard Normal curve, using Table A and the fact that the total area under the curve is 1.

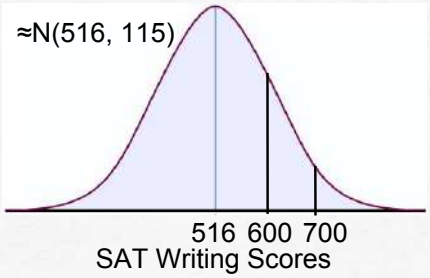
Step 4: *Conclusion.* Write your conclusion in the context of the problem.



Practice Using Z-Table

·An SAT Writing Scores are N(516, 115) What percent are between 600 and 700?

- Step 1 – Sketch the curve, label mean and scores of interest



- Step 2 – Find Zscores
- Step 3 – Use Table to find Area
- Step 4 – Answer in context

APPENDIX - Answers

Answers Sample Test Scores

79	81	80	77	73	83	74	93	78	80	75	67	73
77	83	86	90	79	85	83	89	84	82	77	72	

Julia: z=(86-80)/6.07
z= 0.99 {above average = +z}

Kevin: z=(72-80)/6.07
z= -1.32 {below average = -z}

Katie: z=(80-80)/6.07
z= 0 {average z = 0}

Answers Comparing Scores

- So which test did she perform better?

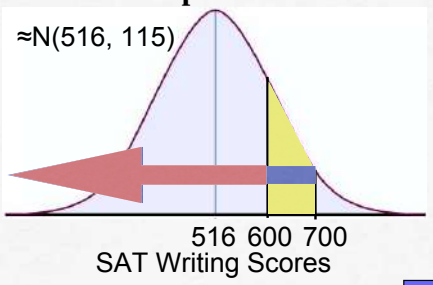
Statistics
$$z = \frac{86 - 80}{6.07} = 0.99$$

Chemistry
$$z = \frac{82 - 76}{4} = 1.5$$

Although she had a lower score, she performed relatively better in Chemistry.

Practice Using Z-Table

- **SAT Writing Scores are N(516, 115)**
  - **What percent are between 600 and 700?**



$$z_{700} = \frac{700 - 516}{115} = \frac{184}{115} = 1.6$$

$$z_{600} = \frac{600 - 516}{115} = \frac{84}{115} = 0.73$$

%Below 700≈.9452 %Below 600≈.7673

%Between 600 and 700≈.9452-.7673≈.1779

4N HW - TRY THESE

4N HW - TRY THESE

Normal distribution, z-score transformation and the use of normal table

6.34 Green Sea Urchins. From the paper "Effects of Chronic Nitrate Exposure on Gonad Growth in Green Sea Urchin *Strongylocentrotus droebachiensis*" (*Aquaculture*, Vol. 242, No. 1–4, pp. 357–363) by S. Siikavuopio et al., we found that weights of adult green sea urchins are normally distributed with mean 52.0 g and standard deviation 17.2 g. Let x denote weight of adult green sea urchins.

a. Sketch the distribution of the variable x.
b. Obtain the standardized version, z, of x.
c. Identify and sketch the distribution of z.
d. The percentage of adult green sea urchins with weights between 50 g and 60 g is equal to the area under the standard normal curve between _____ and _____.
e. The percentage of adult green sea urchins with weights above 40 g is equal to the area under the standard normal curve that lies to the _____ of _____.

Use Green Tables to obtain the areas under the standard normal curve required in Exercises 6.55– 6.62.

Sketch a standard normal curve and shade the area of interest in each problem.

6.56 Determine the area under the standard normal curve that lies to the left of
a. −0.87. b. 3.56. c. 5.12.

6.58 Find the area under the standard normal curve that lies to the right of
a. 2.02. b. −0.56. c. −4.

6.60 Determine the area under the standard normal curve that lies between
a. −0.88 and 2.24. b. −2.5 and −2.
c. 1.48 and 2.72. d. −5.1 and 1.

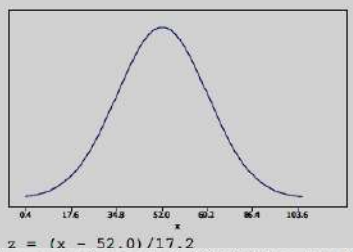
6.62 Find the area under the standard normal curve that lies
a. either to the left of −1 or to the right of 2.
b. either to the left of −2.51 or to the right of −1.

4N HW

TRY THESE

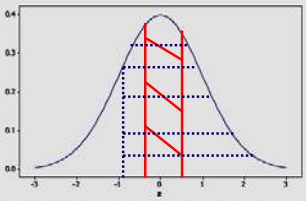
6.34

(a)



(b) z = (x − 52.0)/17.2

(c) z has a standard normal distribution (μ = 0 and σ = 1).



(d) The percentage of adult green sea urchins with weights between 50 g and 60 g is equal to the area under the standard normal curve between −0.12 and 0.47.

(e) The percentage of adult green sea urchins with weights above 40 g is equal to the area under the standard normal curve that lies to the right of −0.70.

(d) Solid lines
$z_{-.12} = .4522$ $z_{.47} = .6808$
Area=.6808-.4522=.2286

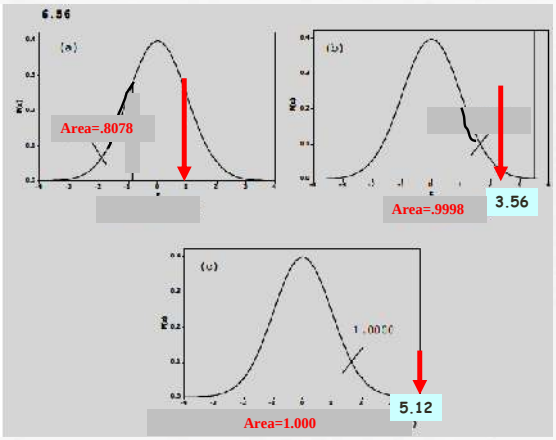
(e) Dotted lines
$z_{-.70} = .2420$
Area= 1-.2420=.7580

Area to the LEFT of Z-Scores

(comes directly from the Z-Table)

6.56

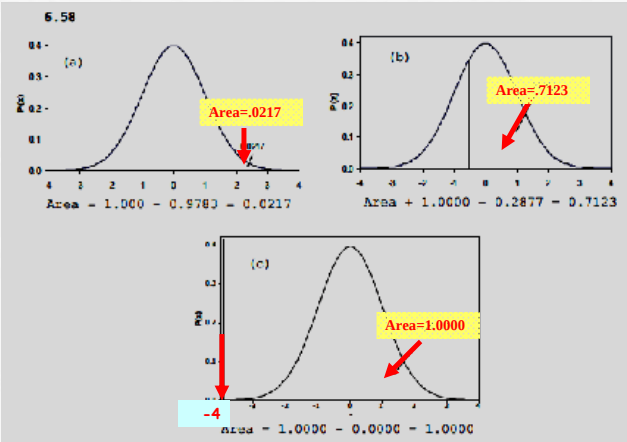
6.56



Area to the RIGHT of Z-Scores

(1.000 minus probability from the Z-Table)

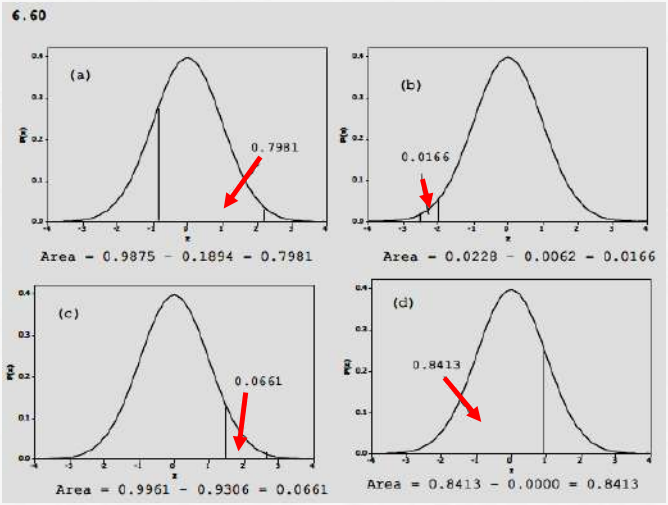
6.58



Area BETWEEN the Z-Scores

(Subtract probabilities)

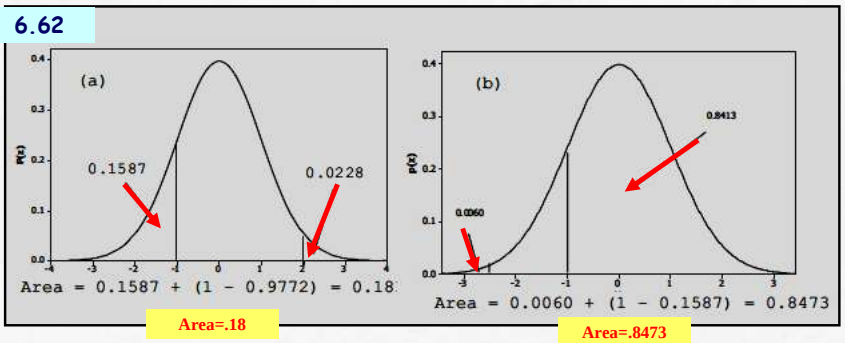
6.60



Area at the tails of the Z-Scores

(Add probabilities)

6.62



Area=.18

Area=.8473