Lecture 17: Chapter 7, Section 3 Continuous Random Variables; Normal Distribution

- Relevance of Normal Distribution
- Continuous Random Variables
- 68-95-99.7 Rule for Normal R.V.s
- Standardizing/Unstandardizing
- Probabilities for Standard/Non-standard Normal R.V.s

Looking Back: *Review*

- **4 Stages of Statistics**
  - Data Production (discussed in Lectures 1-4)
  - Displaying and Summarizing (Lectures 5-12)
  - Probability
    - Finding Probabilities (discussed in Lectures 13-14)
    - Random Variables (introduced in Lecture 15)
      - Binomial (discussed in Lecture 16)
      - Normal
    - Sampling Distributions
  - Statistical Inference

Role of Normal Distribution in Inference

- **Goal:** Perform inference about unknown population proportion, based on sample proportion
- **Strategy:** Determine behavior of sample proportion in random samples with known population proportion
- **Key Result:** Sample proportion follows normal curve for large enough samples.

***Looking Ahead:** Similar approach will be taken with means.*

Discrete vs. Continuous Distributions

- **Binomial Count X**
  - **discrete** (distinct possible values like numbers 1, 2, 3, …)
- **Sample Proportion** $\hat{p} = \frac{X}{n}$
  - **also discrete** (distinct values like count)
- **Normal Approx. to Sample Proportion**
  - **continuous** (follows normal curve)
  - Mean p, standard deviation $\sqrt{\frac{p(1-p)}{n}}$

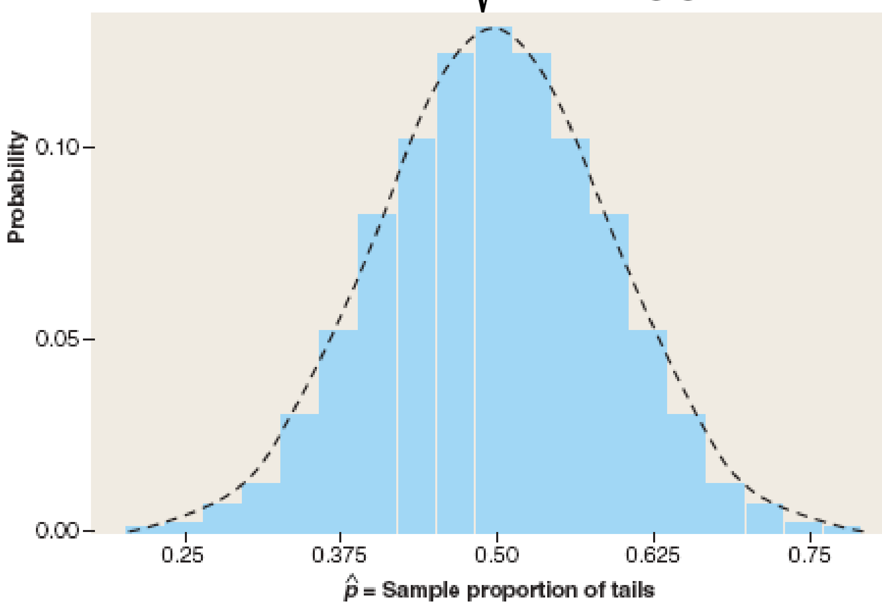
Sample Proportions Approx. Normal *(Review)*

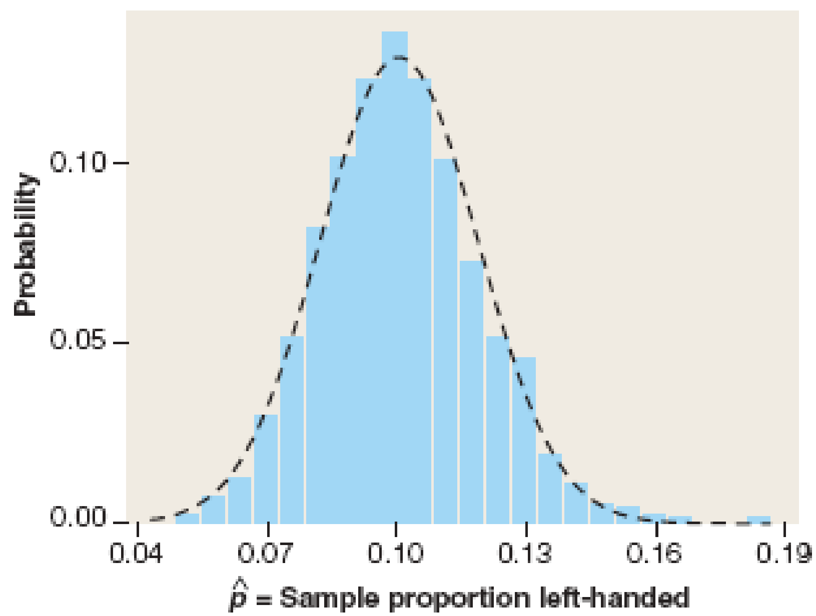
- Proportion of tails in n=16 coinflips (p=0.5) has

$\mu = 0.5, \sigma = \sqrt{\frac{0.5(1-0.5)}{16}} = 0.125$, shape approx normal

- Proportion of lefties (p=0.1) in n=100 people has

$\mu = 0.1, \sigma = \sqrt{\frac{0.1(1-0.1)}{100}} = 0.03$, shape approx normal





Example: *Variable Types*

- **Background**: Variables in survey excerpt:

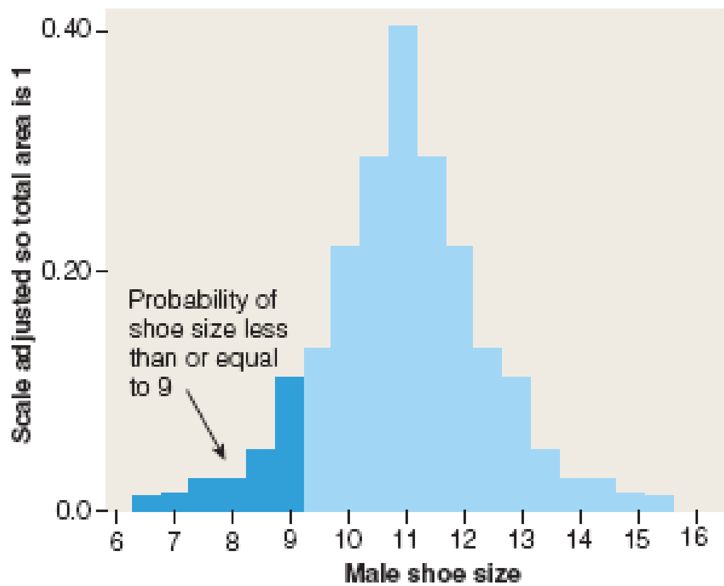
age	breakfast?	comp	credits	...
19.67	no	120	15	
20.08	no	120	16	
19.08	yes	40	14	
...	...	...	...	

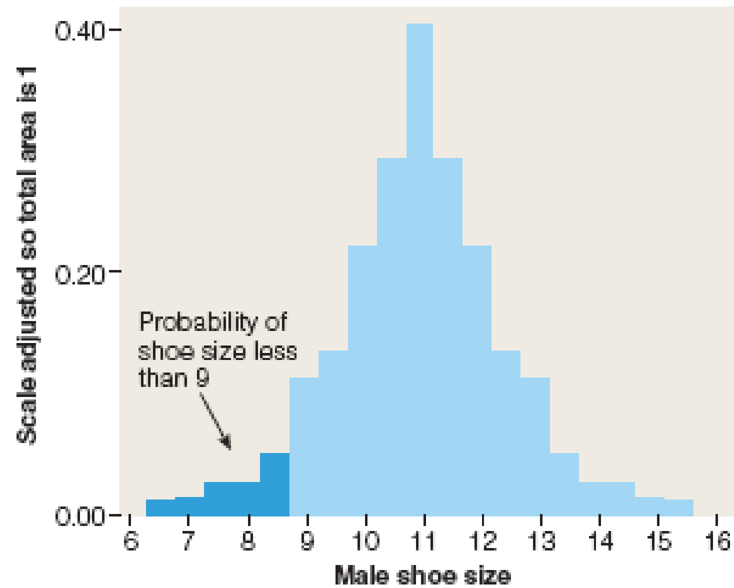
- **Question:** Identify type (cat, discrete quan, continuous quan)
  - Age? Breakfast? Comp (daily min. on computer)? Credits?
- **Response:**
  - Age:
  - Breakfast:
  - Comp (daily time in min. on computer)
  - Credits:

Probability Histogram for Discrete R.V.

Histogram for male shoe size X represents probability by area of bars

- $P(X \leq 9)$ (on left)
- $P(X < 9)$ (on right)



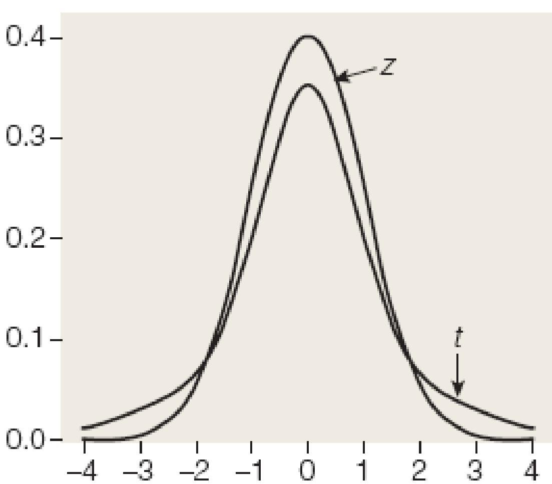


For discrete R.V., strict inequality or not matters.

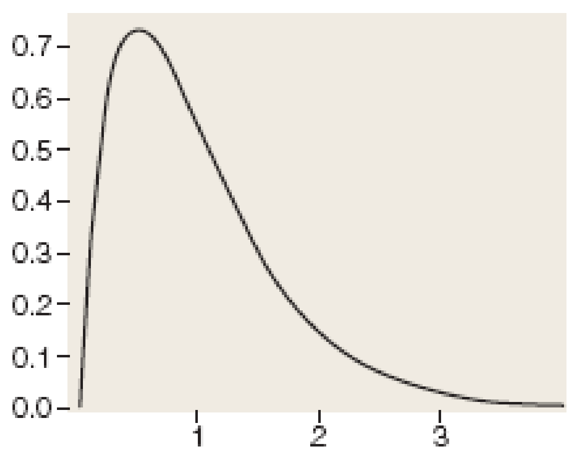
Definition

Density curve: smooth curve showing prob. dist. of continuous R.V. Area under curve shows prob. that R.V. takes value in given interval.

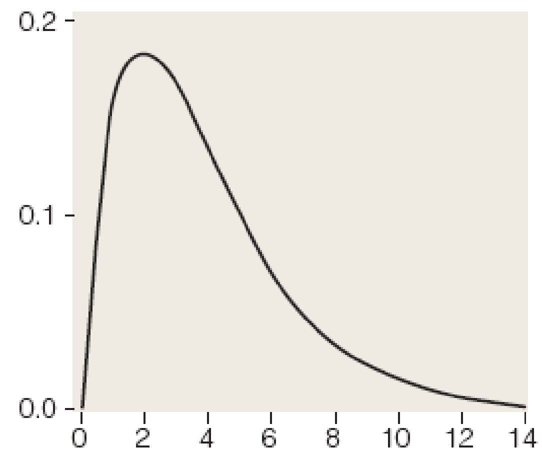
Looking Ahead: *Most commonly used density curve is* ***normal z*** *but to perform inference we also use* ***t, F, and chi-square*** *curves.*



z or *t* = Standardized difference between sample mean and proposed population mean



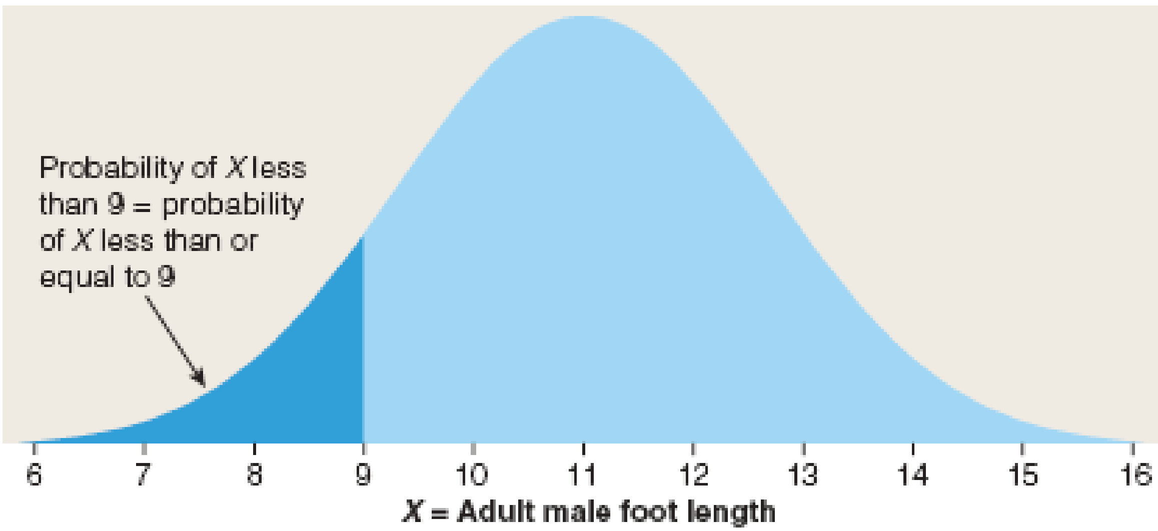
F = Standardized difference among observed sample means



chi-square = Standardized difference between counts observed and counts expected if proportions in various responses were equal for various explanatory groups (in 3 by 3 table)

Density Curve for Continuous R.V.

Density curve for male foot length X represents probability by area under curve.



$$P(X \leq 9) = P(X < 9)$$

Continuous RV: strict inequality or not doesn't matter.

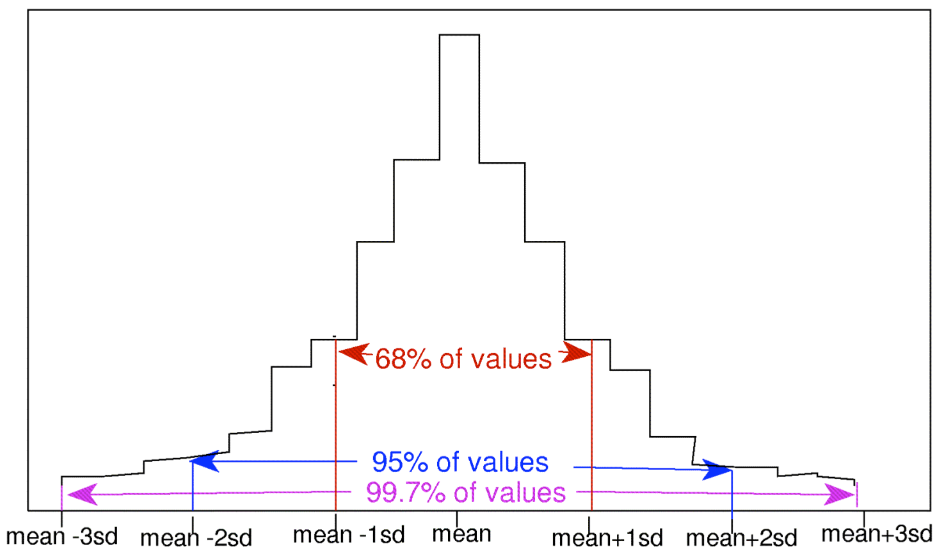
***A Closer Look:** Shoe sizes are discrete; foot lengths are continuous.*

68-95-99.7 Rule for Normal Data *(Review)*

Values of a normal **data set** have

- 68% within 1 standard deviation of mean
- 95% within 2 standard deviations of mean
- 99.7% within 3 standard deviations of mean

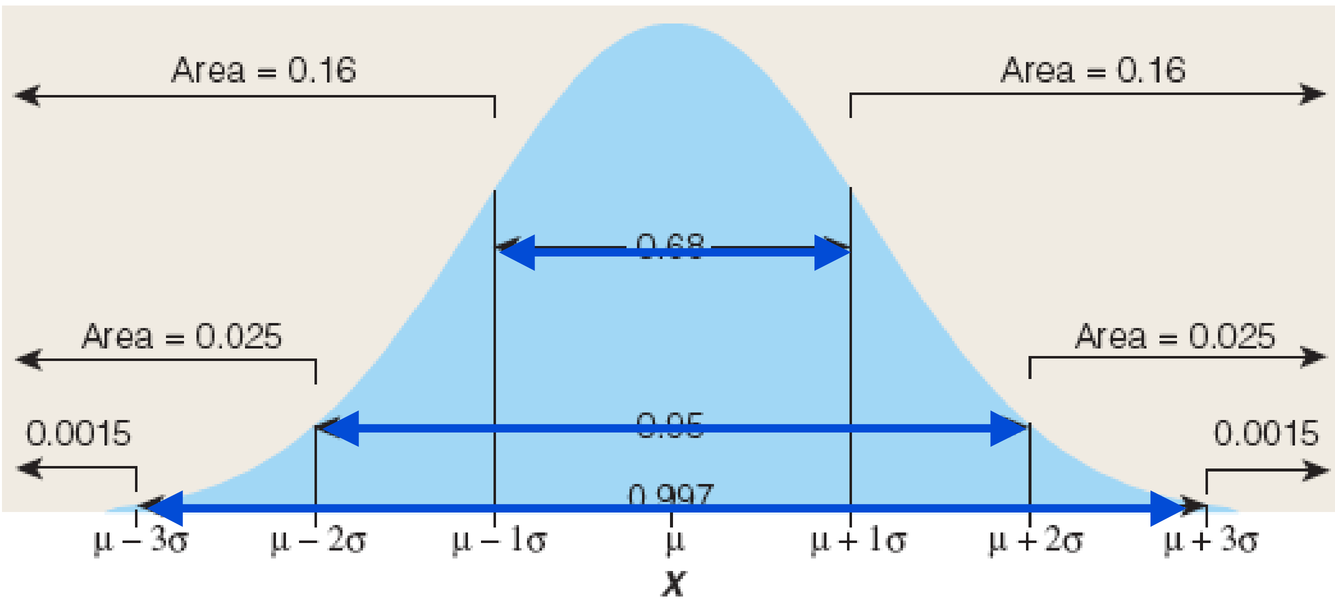
68-95-99.7 Rule for Normal Distributions



68-95-99.7 Rule: Normal Random Variable

Sample at **random** from normal **population**; for sampled value *X* (a R.V.), probability is

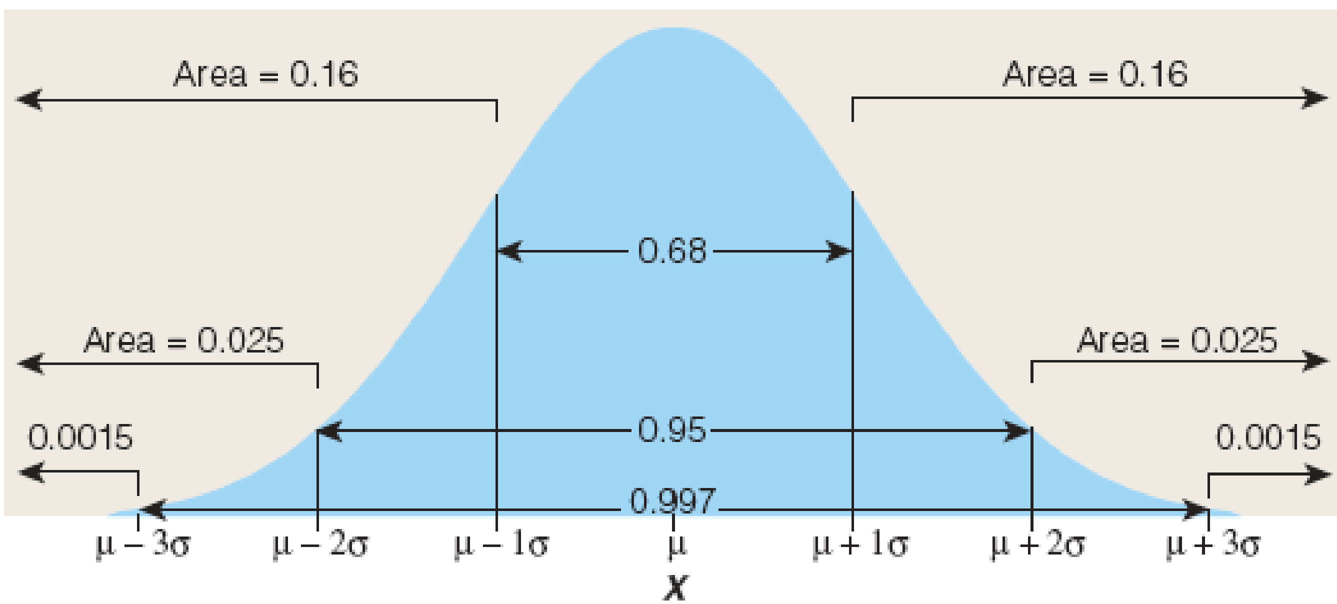
- 68% that *X* is within 1 standard deviation of mean
- 95% that *X* is within 2 standard deviations of mean
- 99.7% that *X* is within 3 standard deviations of mean



68-95-99.7 Rule: Normal Random Variable

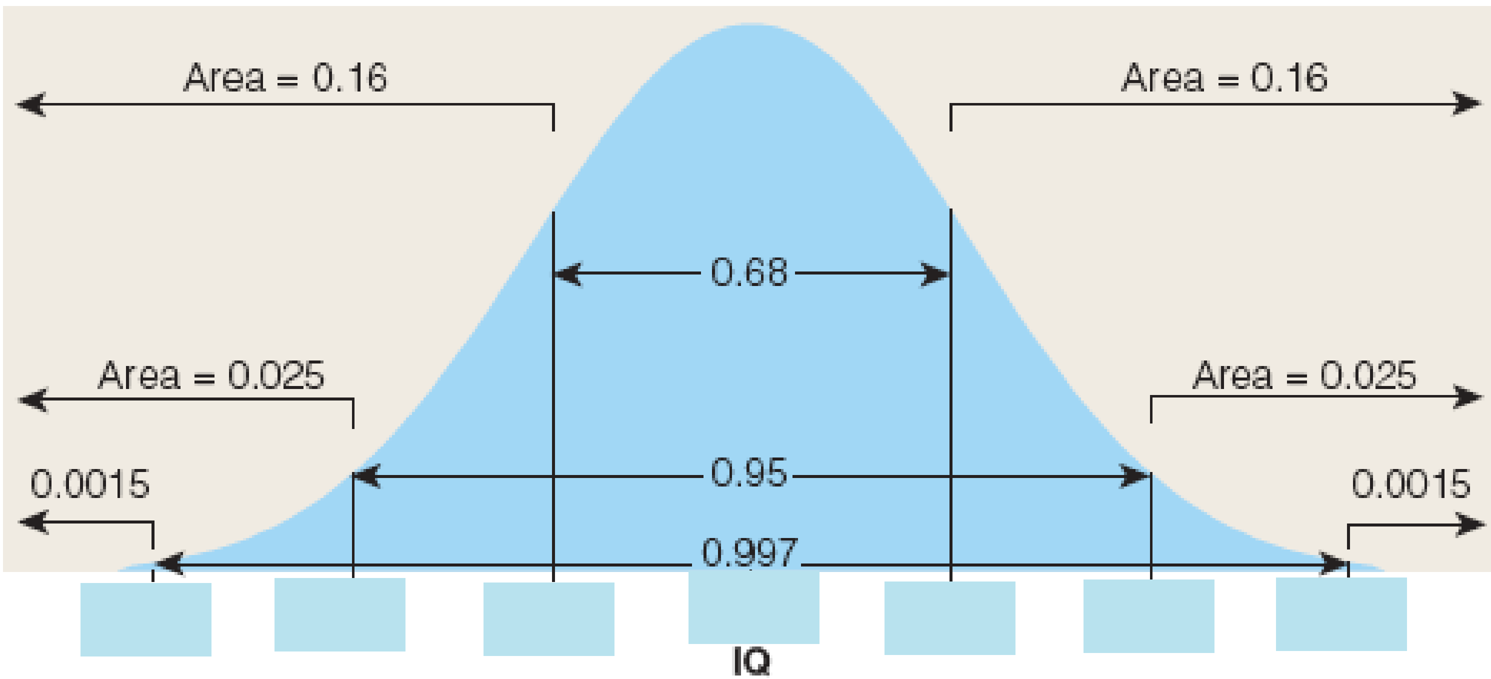
Looking Back: *We use Greek letters to denote population mean and standard deviation.*

mean $= \mu$, standard deviation $= \sigma$



Example: *68-95-99.7 Rule for Normal R.V.*

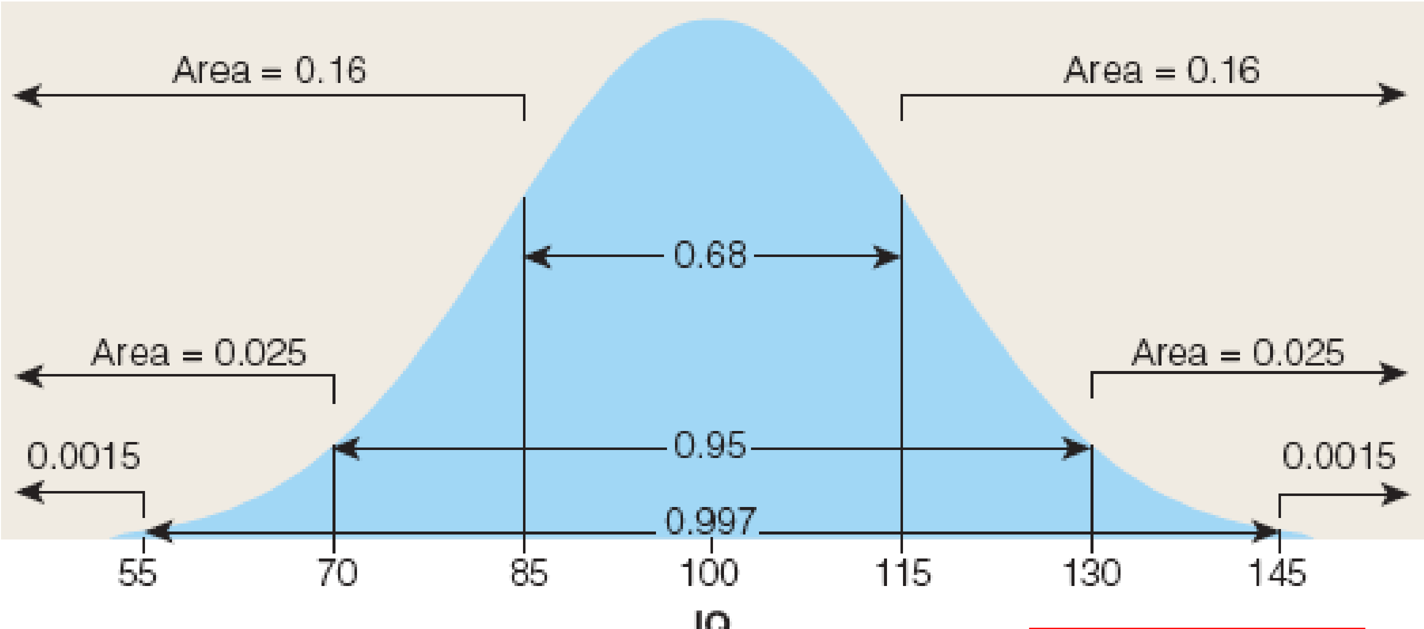
- **Background**: IQ for randomly chosen adult is normal R.V. X with $\mu = 100, \sigma = 15$.
- **Question:** What does Rule tell us about distribution of X?
- **Response:** We can sketch distribution of X:



Practice: 7.47 p.331

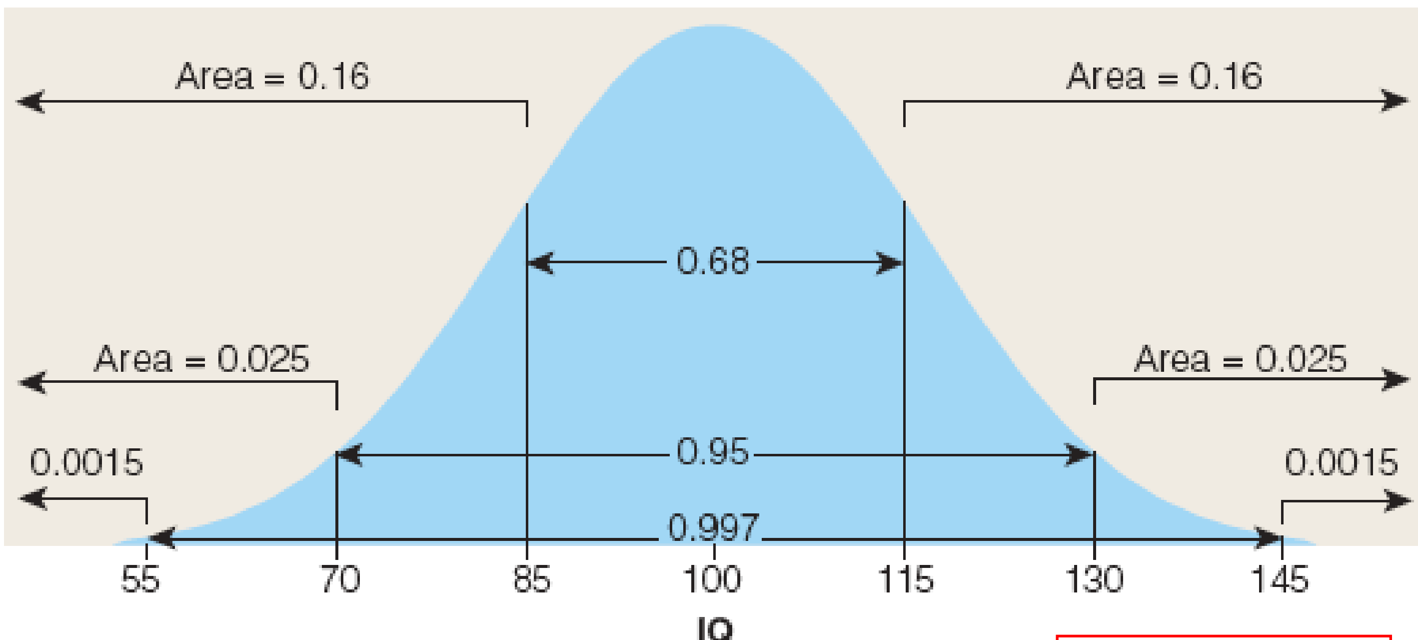
Example: *Finding Probabilities with Rule*

- **Background**: IQ for randomly chosen adult is normal R.V. *X* with $\mu = 100, \sigma = 15.$
- **Question:** Prob. of IQ between 70 and 130 = ?
- **Response:**



Example: *Finding Probabilities with Rule*

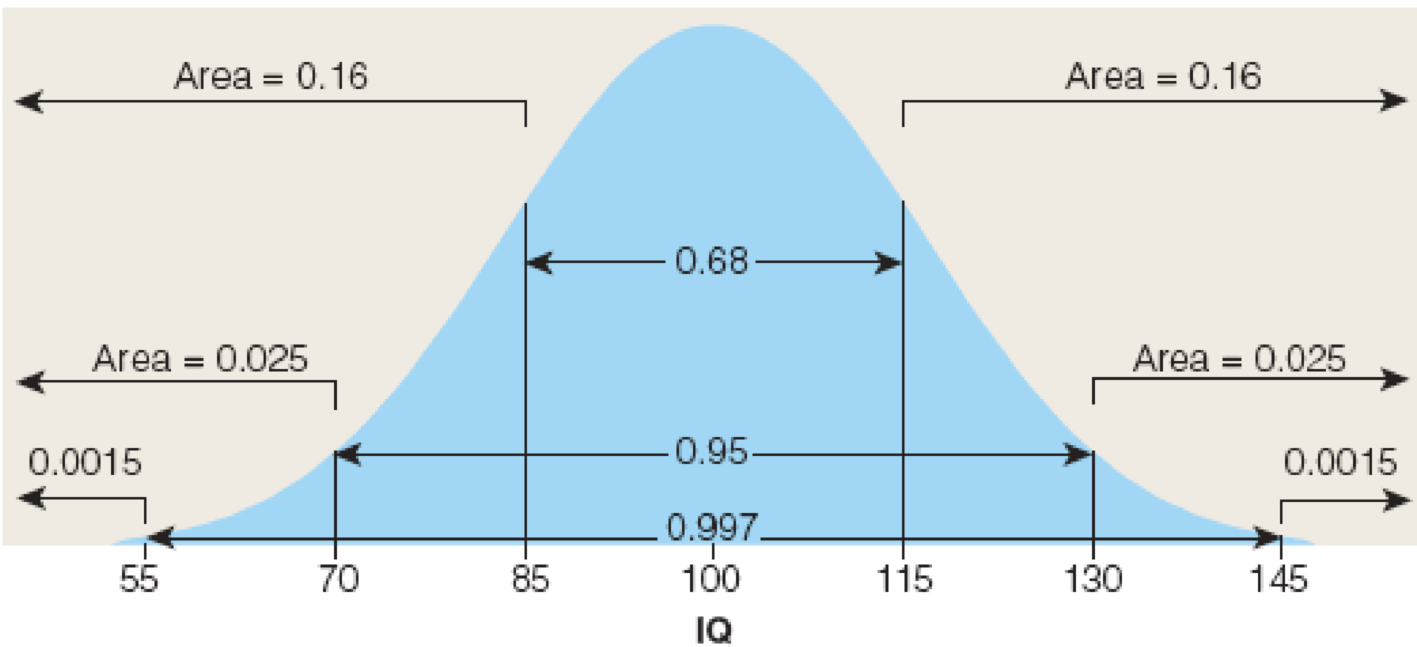
- **Background**: IQ for randomly chosen adult is normal R.V. *X* with $\mu = 100, \sigma = 15.$
- **Question:** Prob. of IQ less than 70 = ?
- **Response:**



Practice: 7.49a p.331

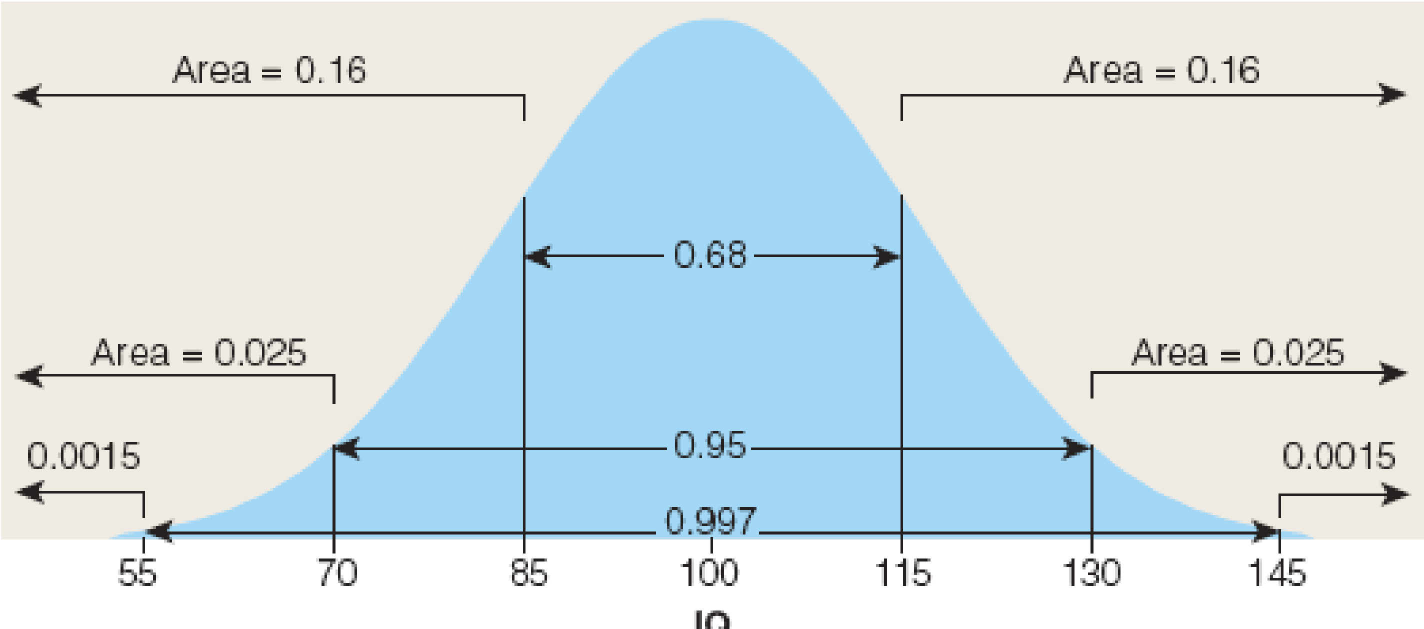
Example: *Finding Probabilities with Rule*

- **Background**: IQ for randomly chosen adult is normal R.V. X with $\mu = 100, \sigma = 15.$
- **Question:** Prob. of IQ less than 100 = ?
- **Response:**



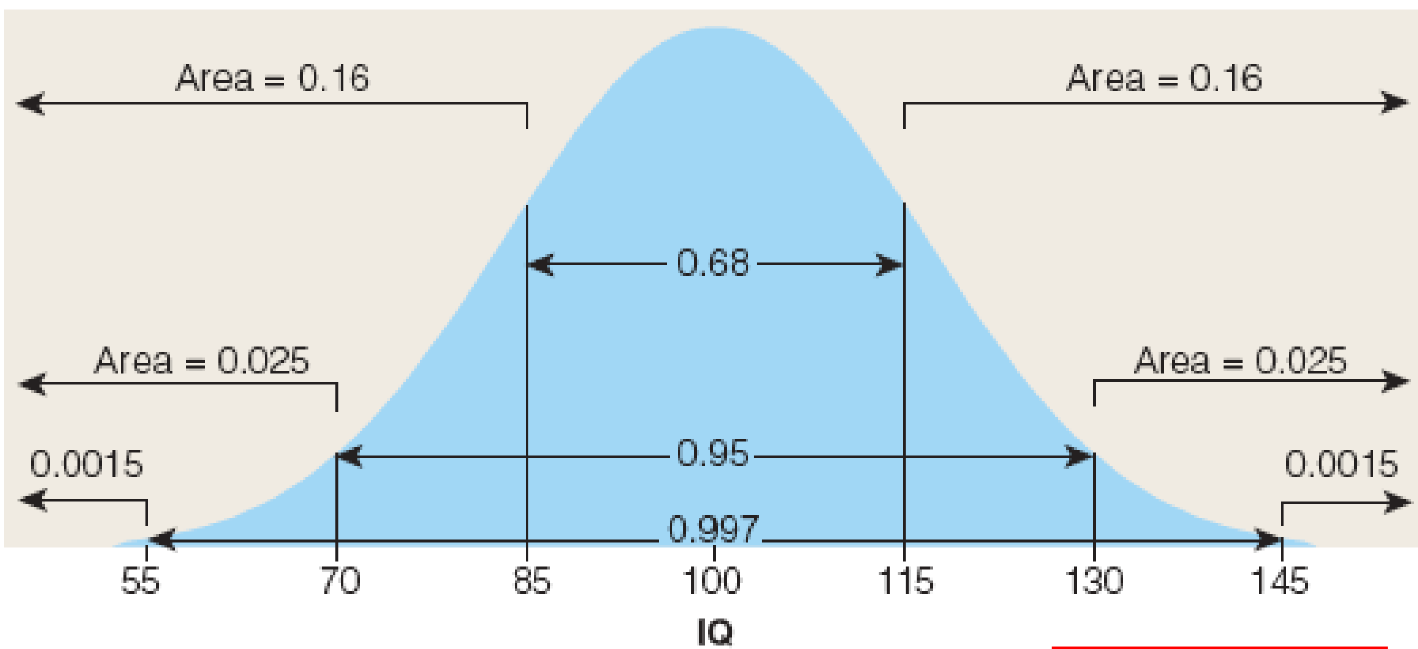
Example: *Finding Values of X with Rule*

- **Background**: IQ for randomly chosen adult is normal R.V. *X* with $\mu = 100$, $\sigma = 15$.
- **Question:** Prob. is 0.997 that IQ is between…?
- **Response:**



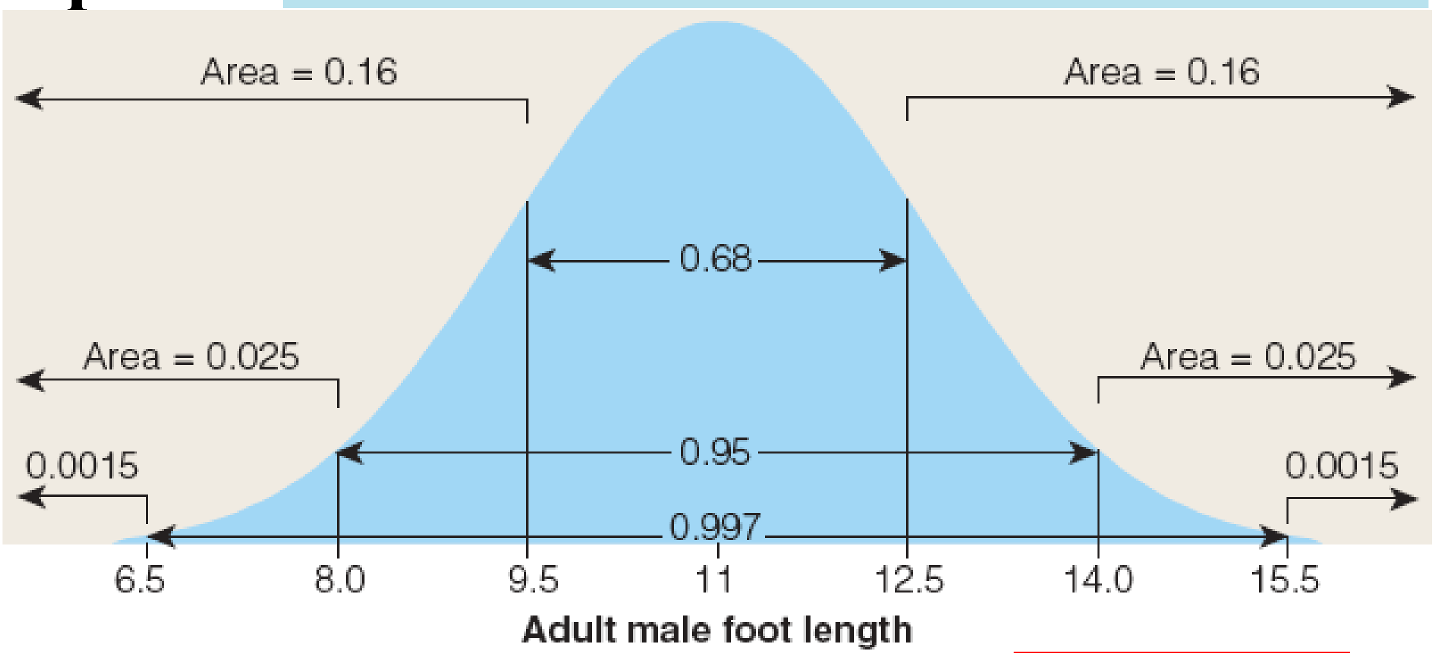
Example: *Finding Values of X with Rule*

- **Background**: IQ for randomly chosen adult is normal R.V. *X* with $\mu = 100, \sigma = 15.$
- **Question:** Prob. is 0.025 that IQ is above…?
- **Response:**



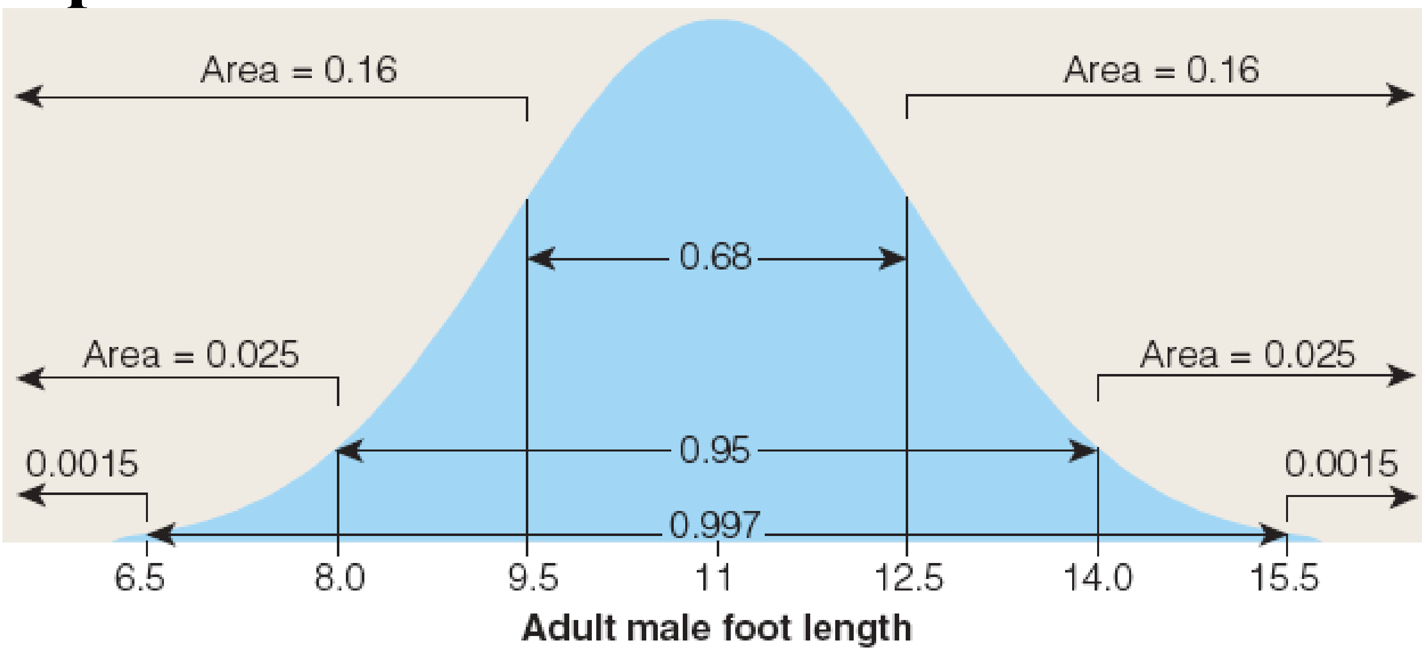
Example: *Using Rule to Evaluate Probabilities*

- **Background**: Foot length of randomly chosen adult male is normal R.V. *X* with $\mu = 11, \sigma = 1.5$ (in.)
- **Question:** How unusual is foot less than 6.5 inches?
- **Response:**



Example: *Using Rule to Estimate Probabilities*

- **Background**: Foot length of randomly chosen adult male is normal R.V. X with $\mu = 11, \sigma = 1.5$ (in.)
- **Question:** How unusual is foot more than 13 inches?
- **Response:**



Practice: 7.49c-d p.331

Definition *(Review)*

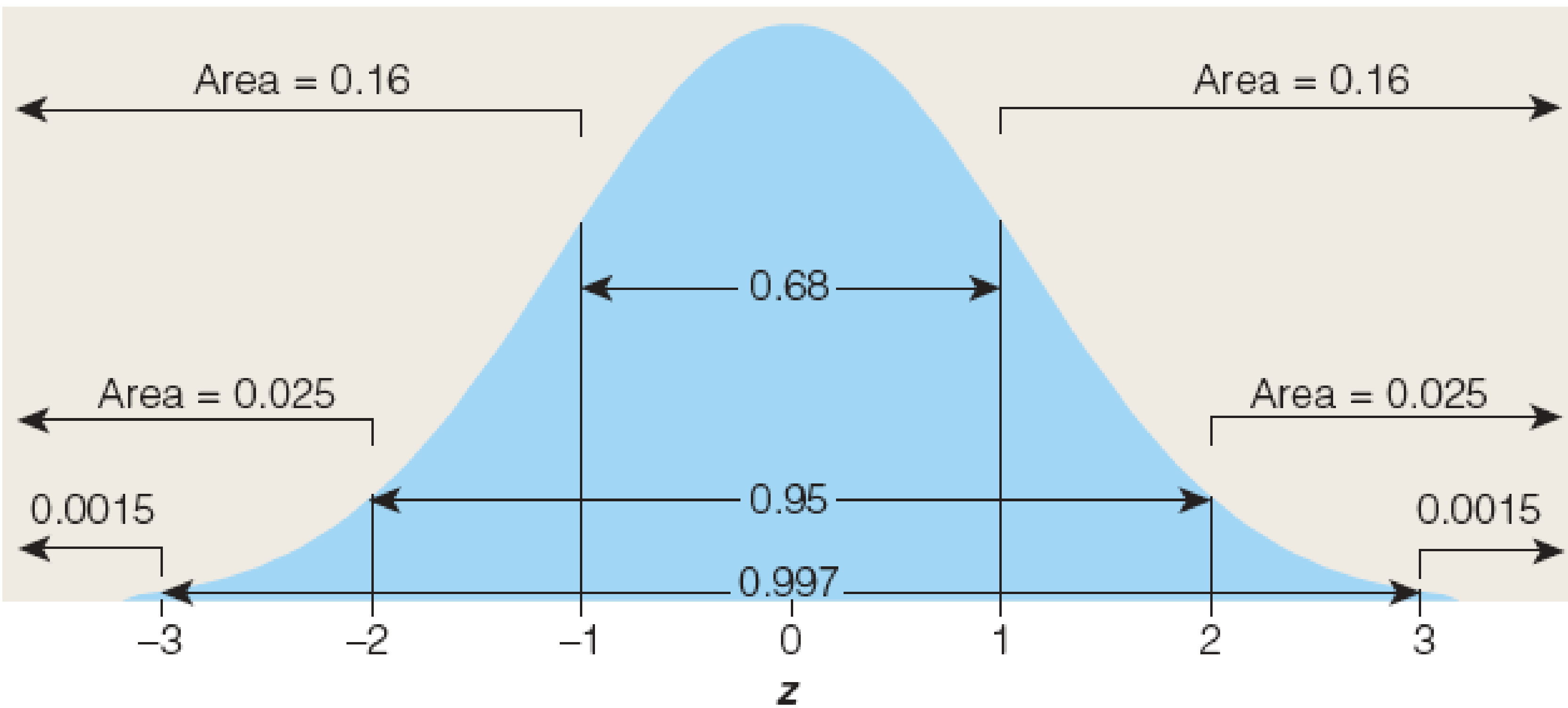
- ***z*-score**, or **standardized value**, tells how many standard deviations below or above the mean the original value is:

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

- Notation for Population: $z = \frac{x - \mu}{\sigma}$
  - $z > 0$ for x above mean
  - $z < 0$ for x below mean
- Unstandardize: $x = \mu + z\sigma$

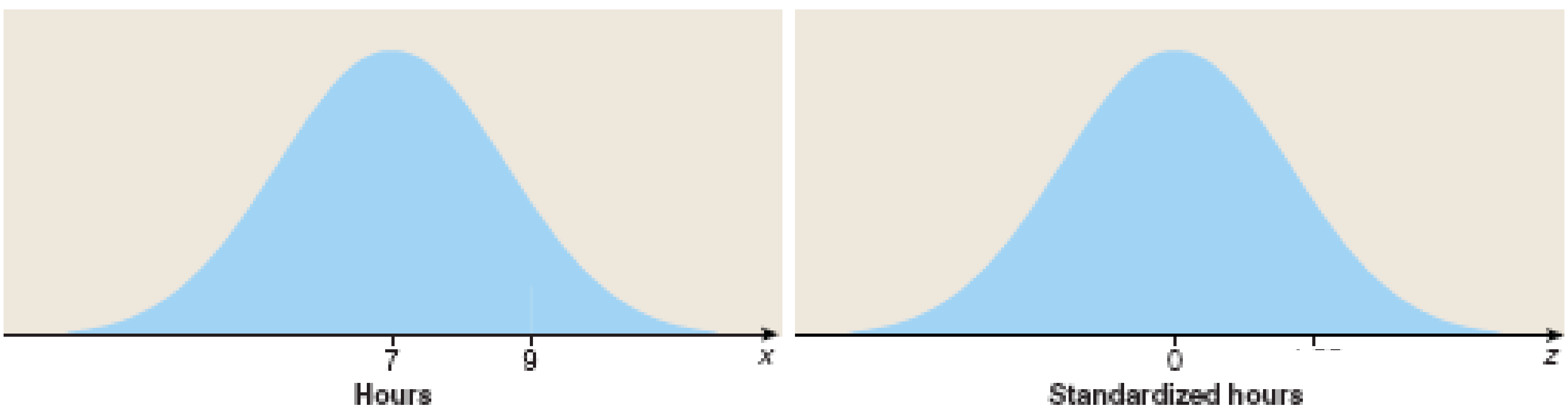
Standardizing Values of Normal R.V.s

Standardizing to *z* lets us avoid sketching a different curve for every normal problem: we can always refer to same standard normal (*z*) curve:



Example: *Standardized Value of Normal R.V.*

- **Background**: Typical nightly hours slept by college students normal; $\mu = 7,\ \sigma = 1.5$
- **Question:** How many standard deviations below or above mean is 9 hours?
- **Response:** Standardize to $z =$
(9 is standard deviations above mean)



Practice: 7.49e p.331

Example: *Standardizing/Unstandardizing Normal R.V.*

- **Background**: Typical nightly hours slept by college students normal; $\mu = 7, \sigma = 1.5.$
- **Questions:**
  - What is standardized value for sleep time 4.5 hours?
  - If standardized sleep time is +2.5, how many hours is it?
- **Responses:**
  - $z =$
  -

Practice: 7.49f-g p.331

Interpreting z-scores *(Review)*

This table classifies ranges of z-scores informally, in terms of being unusual or not.

Size of z	Unusual?
$\|z\|$ greater than 3	extremely unusual
$\|z\|$ between 2 and 3	very unusual
$\|z\|$ between 1.75 and 2	unusual
$\|z\|$ between 1.5 and 1.75	maybe unusual (depends on circumstances)
$\|z\|$ between 1 and 1.5	somewhat low/high, but not unusual
$\|z\|$ less than 1	quite common

Looking Ahead: *Inference conclusions will hinge on whether or not a standardized score can be considered "unusual".*

Example: *Characterizing Normal Values Based on z-Scores*

- **Background**: Typical nightly hours slept by college students normal; $\mu = 7$, $\sigma = 1.5$.
- **Questions:** How unusual is a sleep time of 4.5 hours (z = -1.67)? 10.75 hours (z = +2.5)?
- **Responses:**
  - Sleep time of 4.5 hours (z = -1.67):
  - Sleep time of 10.75 hours (z = +2.5):

Size of z	Unusual?
$\|z\|$ greater than 3	extremely unusual
$\|z\|$ between 2 and 3	very unusual
$\|z\|$ between 1.75 and 2	unusual
$\|z\|$ between 1.5 and 1.75	maybe unusual (depends on circumstances)
$\|z\|$ between 1 and 1.5	somewhat low/high, but not unusual
$\|z\|$ less than 1	quite common

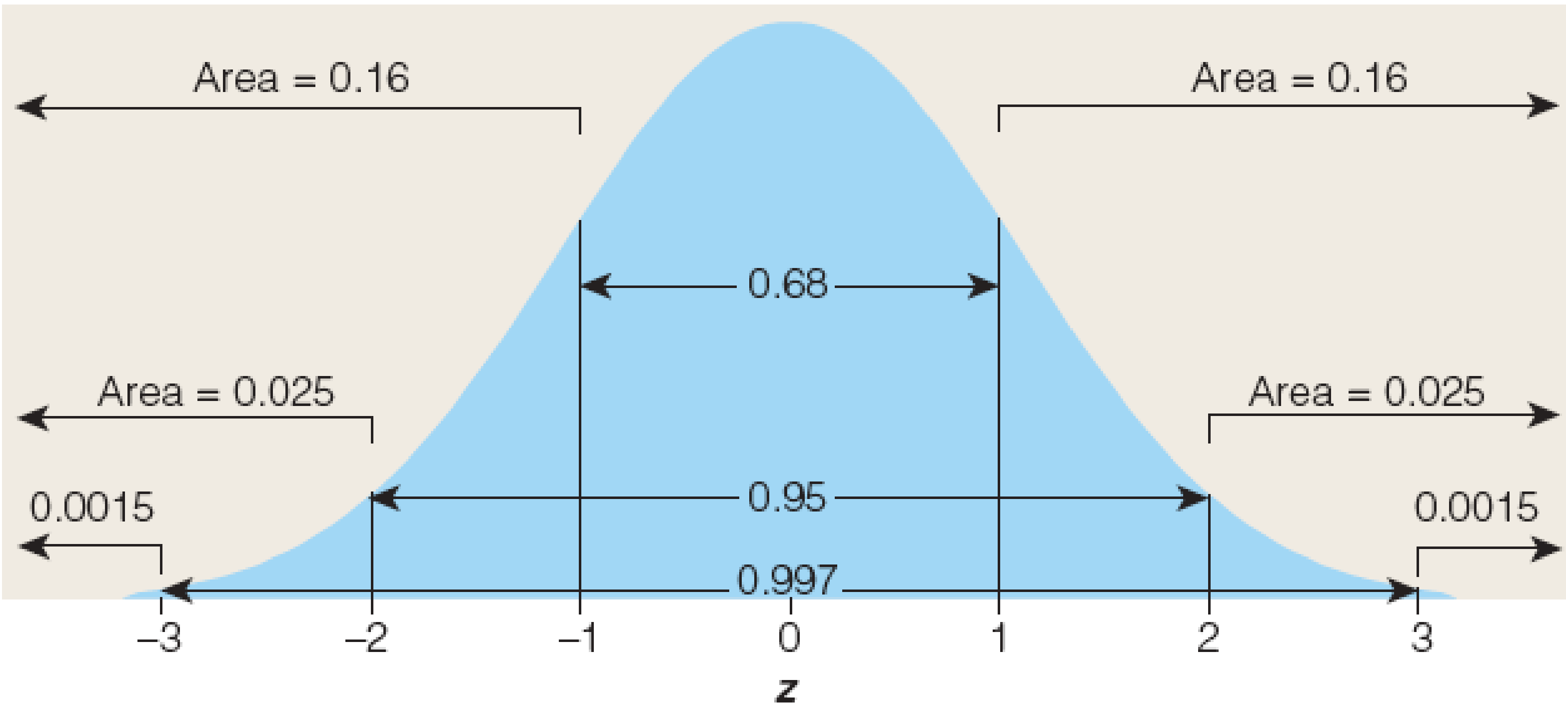
Practice: 7.51 p.332

Normal Probability Problems

- Estimate probability given z
  - Probability close to 0 or 1 for extreme z
- Estimate z given probability
- Estimate probability given non-standard x
- Estimate non-standard x given probability

Example: *Estimating Probability Given z*

- **Background**: Sketch of 68-95-99.7 Rule for *Z*

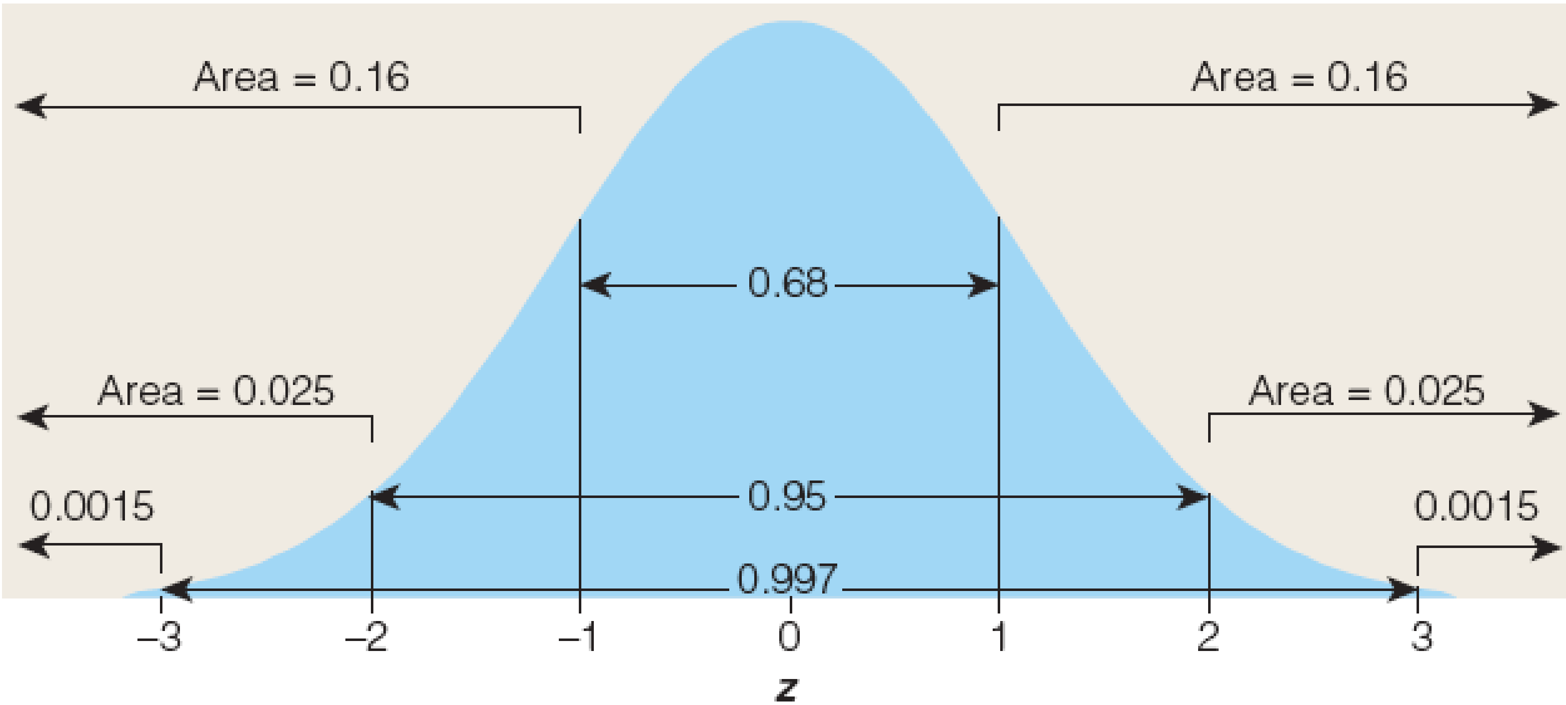


- **Question:** Estimate **P(*Z*<-1.47)**?
- **Response:**

Practice: 7.53 p.332

Example: *Estimating Probability Given z*

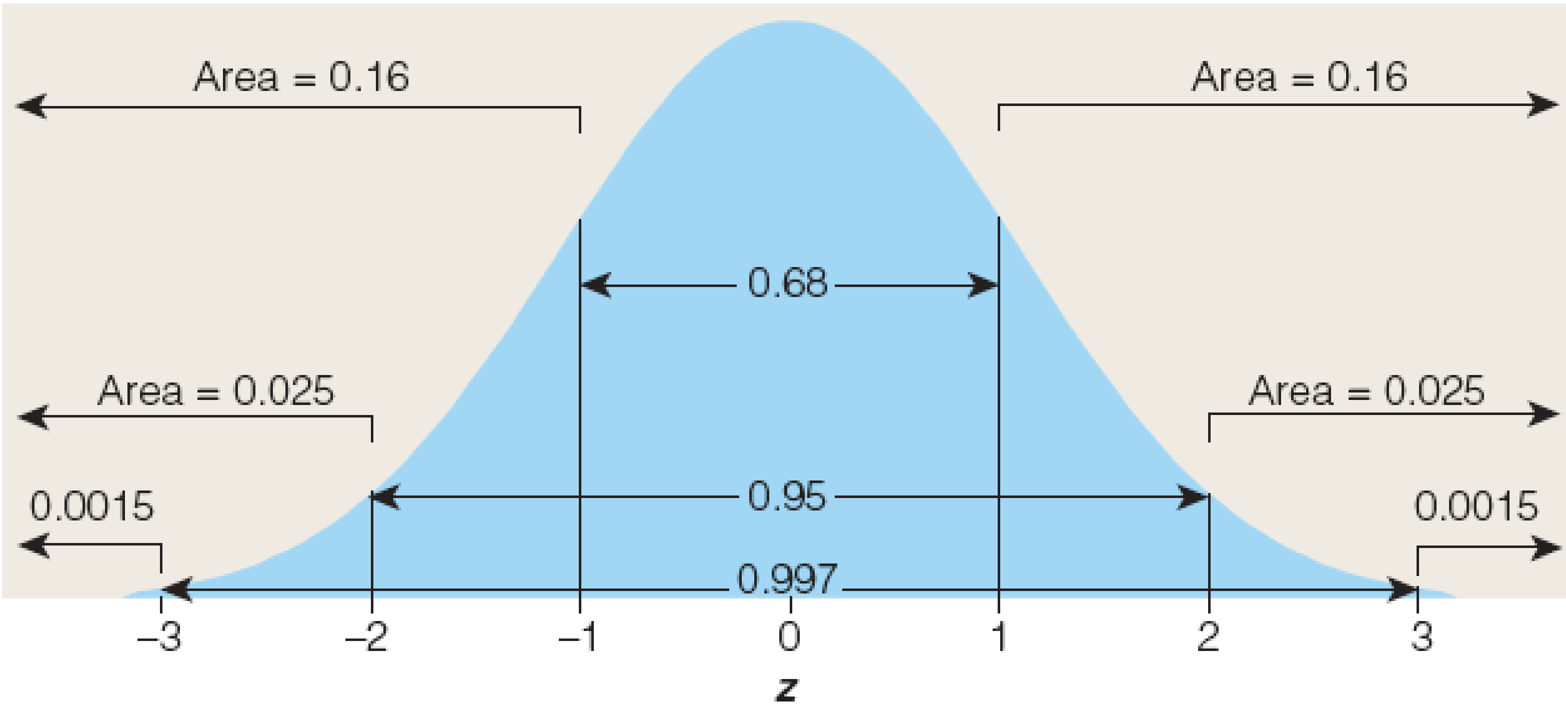
- **Background**: Sketch of 68-95-99.7 Rule for *Z*



- **Question:** Estimate **P(*Z*>+0.75)**?
- **Response:**

Example: *Estimating Probability Given z*

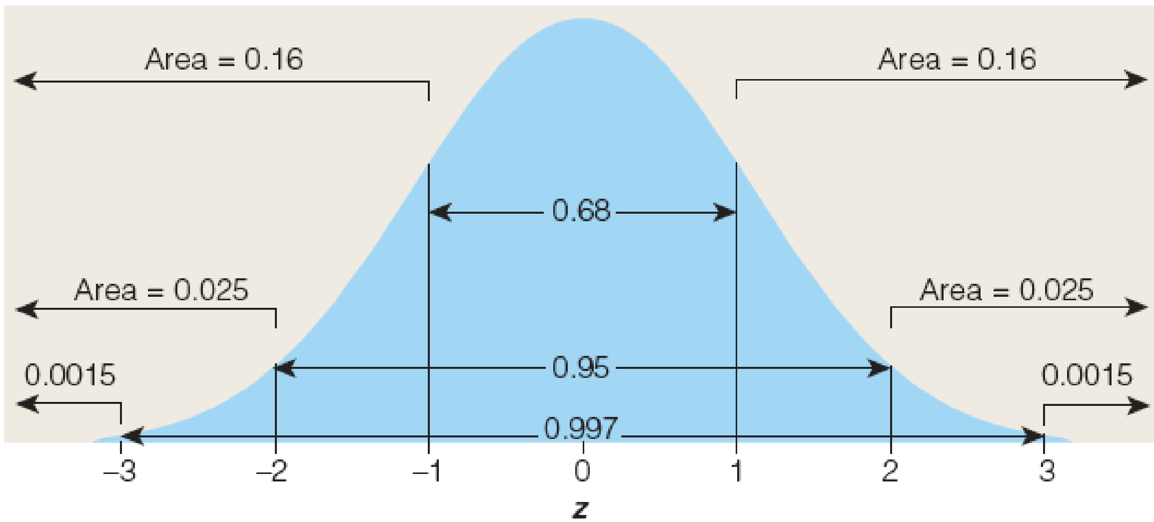
- **Background**: Sketch of 68-95-99.7 Rule for *Z*



- **Question:** Estimate **P(*Z*<+2.8)**?
- **Response:**

Example: *Probabilities for Extreme z*

- **Background**: Sketch of 68-95-99.7 Rule for Z



- **Question:** What are the following (approximately)?

a. $P(Z<-14.5)$ b. $P(Z<+13)$ c. $P(Z>+23.5)$ d. $P(Z>-12.1)$

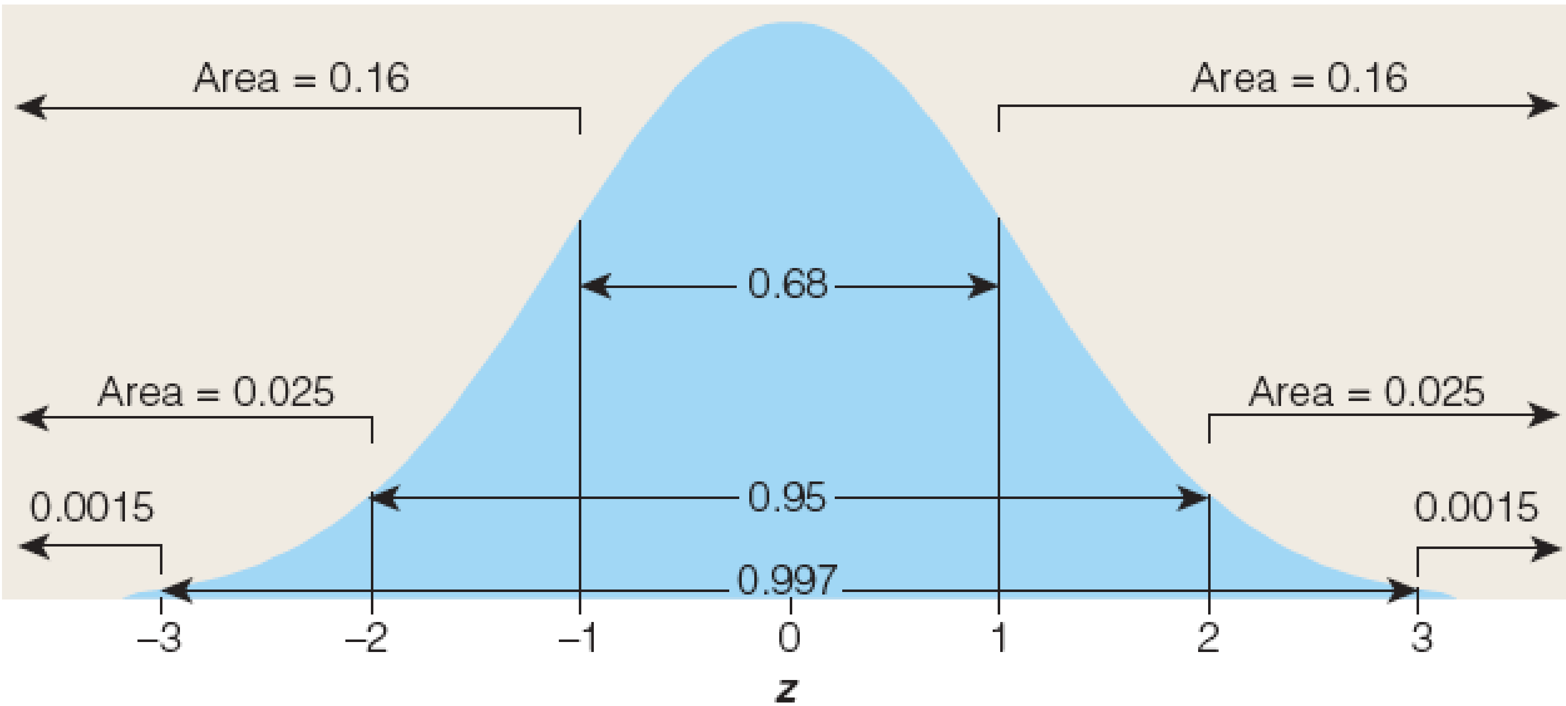
- **Response:**

a. b. c. d.

Practice: 7.54 p.332

Example: *Estimating z Given Probability*

- **Background**: Sketch of 68-95-99.7 Rule for *Z*

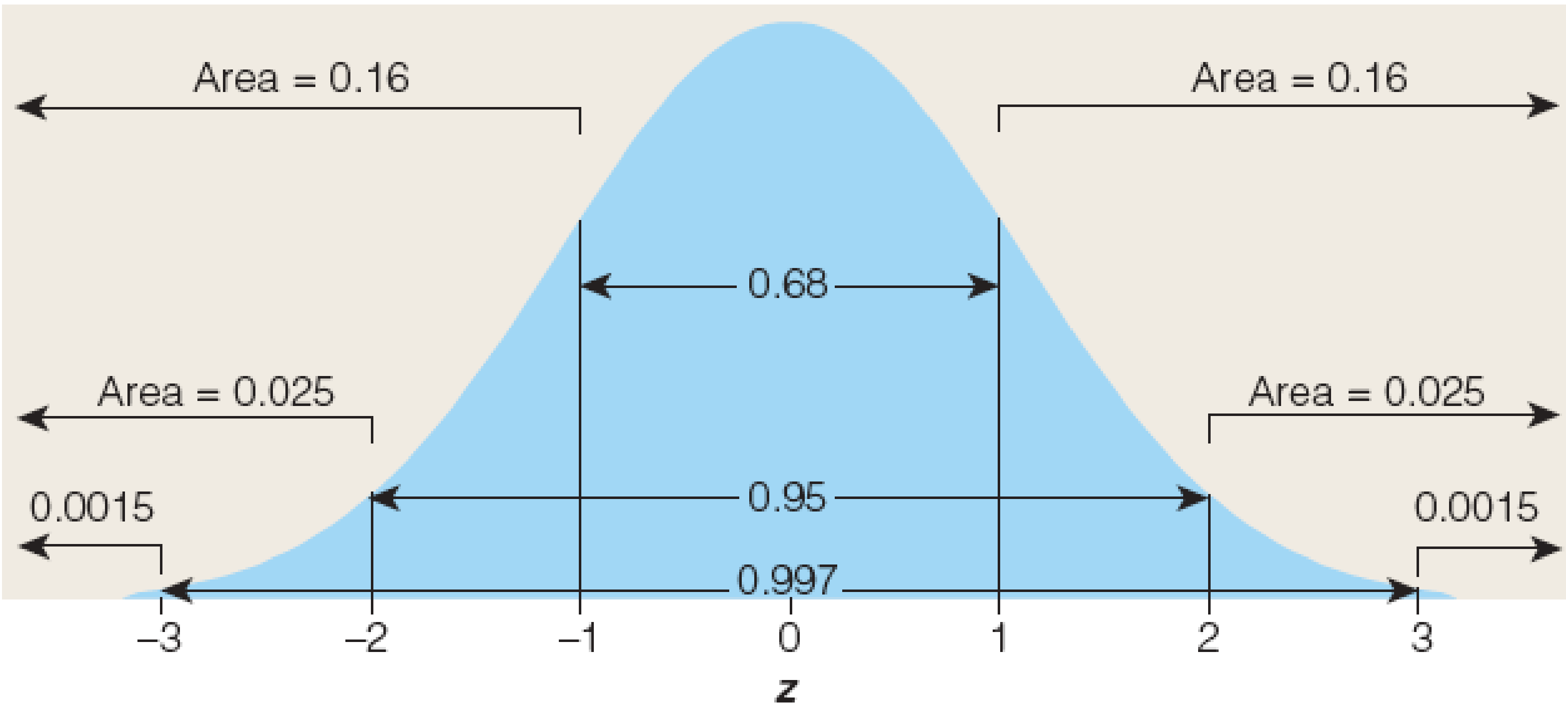


- **Question:** Prob. is 0.01 that $Z <$ what value?
- **Response:**

Practice: 7.57 p.332

Example: *Estimating z Given Probability*

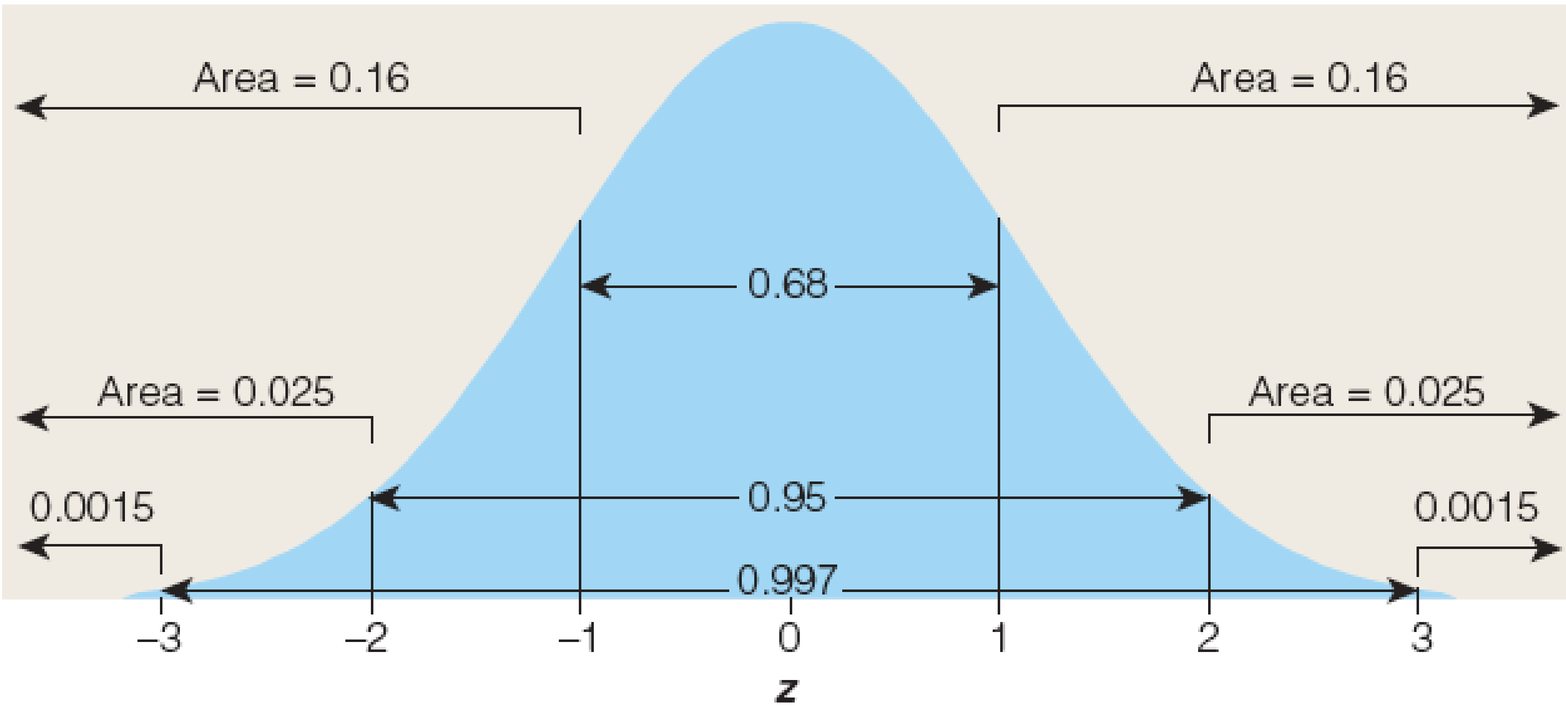
- **Background**: Sketch of 68-95-99.7 Rule for *Z*



- **Question:** Prob. is 0.15 that *Z* >what value?
- **Response:**

Example: *Estimating Probability Given x*

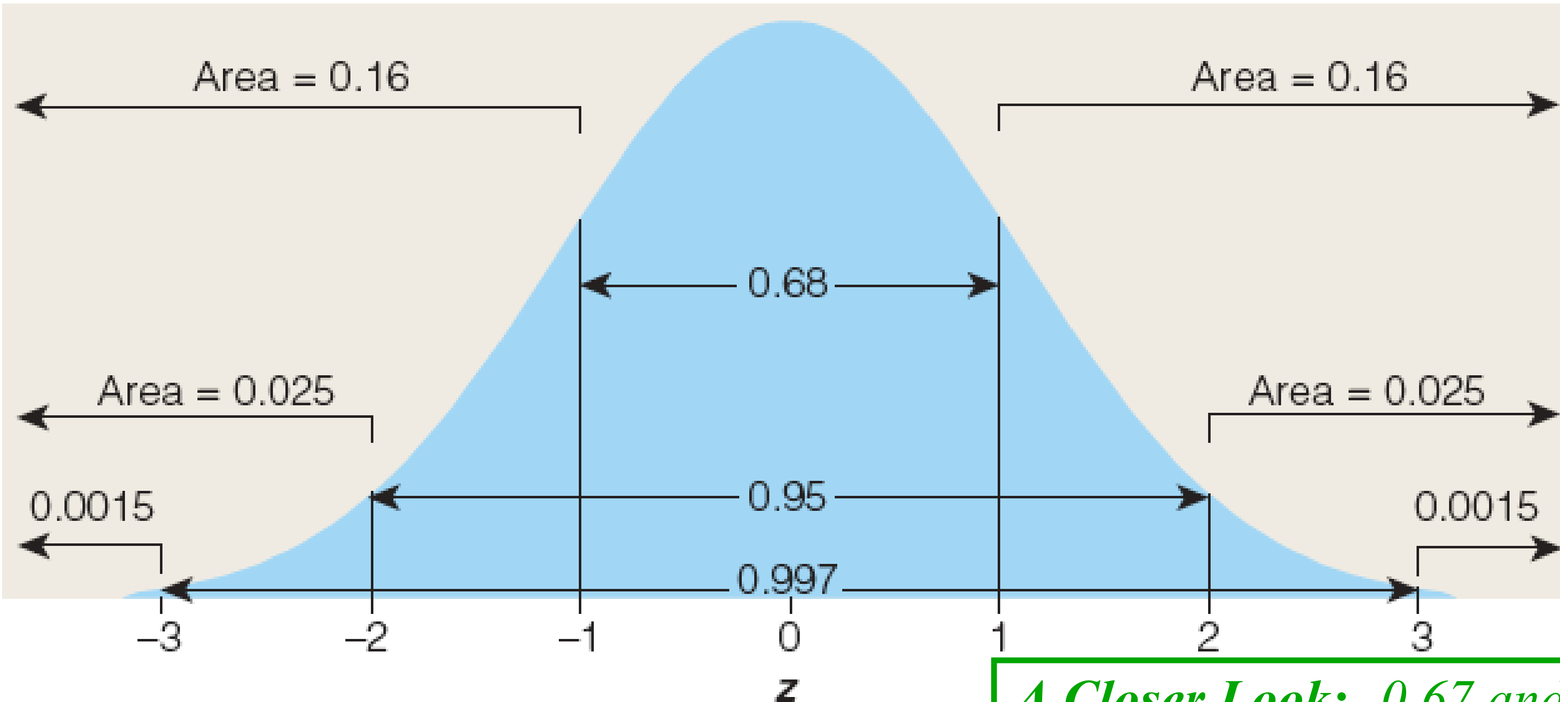
- **Background**: Hrs. slept *X* normal; $\mu = 7,\ \sigma = 1.5.$



- **Question:** Estimate **P(*X*>9)**?
- **Response:**

Example: *Estimating Probability Given x*

- **Background**: Hrs. slept *X* normal; $\mu = 7, \sigma = 1.5.$



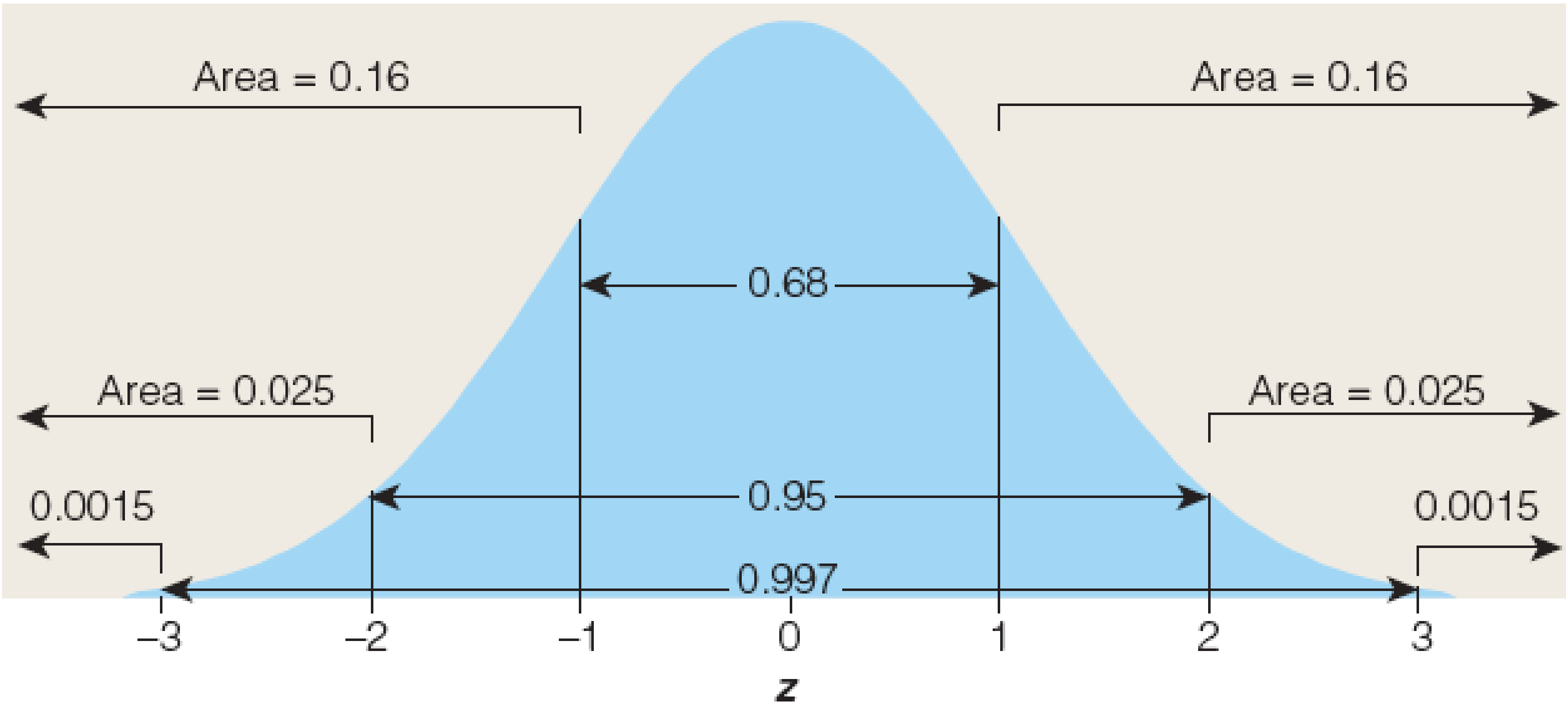
- **Question:** Estimate **P(6<*X*<8)**?

A Closer Look: *-0.67 and +0.67 are the* ***quartiles*** *of the* ***z*** *curve.*

- **Response:**

Example: *Estimating x Given Probability*

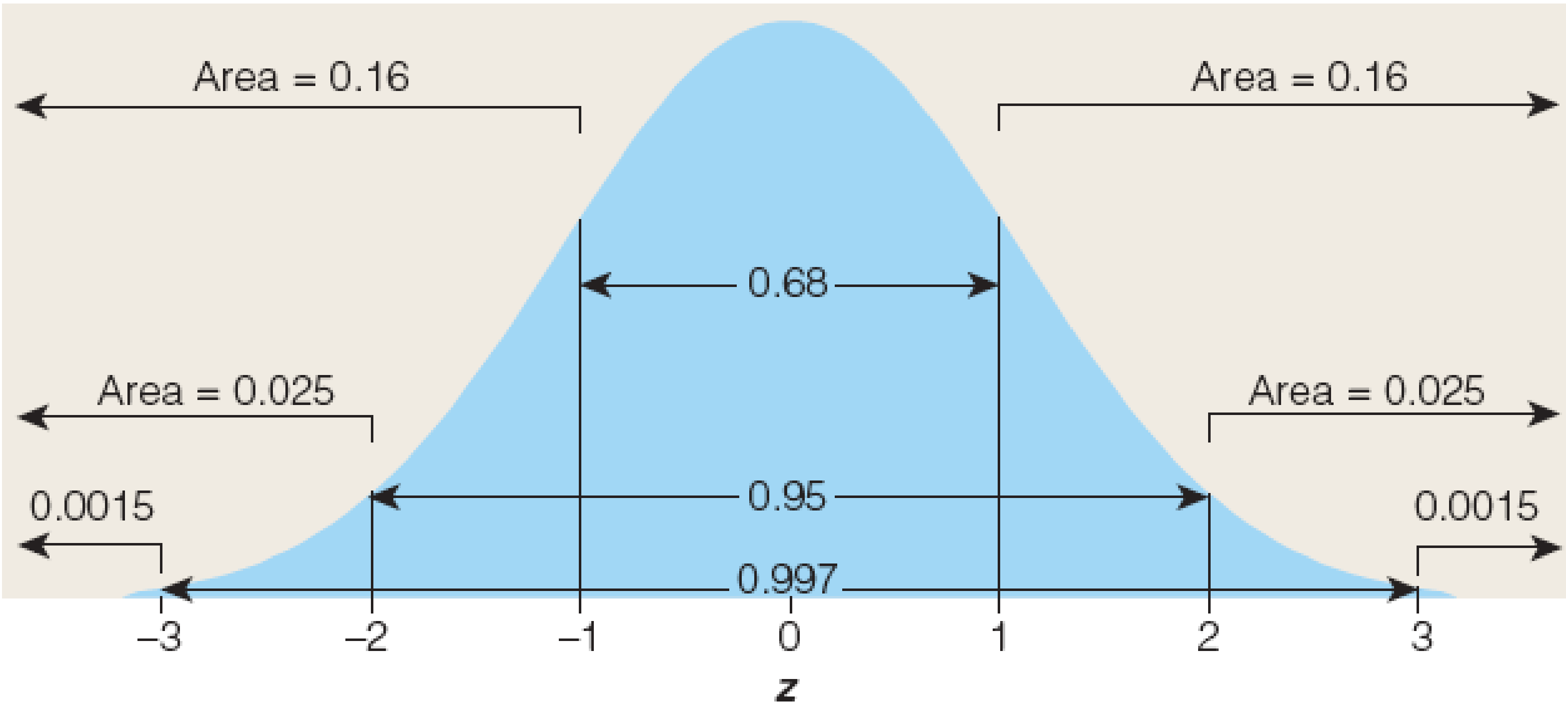
- **Background**: Hrs. slept X normal; $\mu = 7$, $\sigma = 1.5$.



- **Question:** 0.04 is P(X < ?)
- **Response:**

Example: *Estimating x Given Probability*

- **Background**: Hrs. slept X normal; $\mu = 7$, $\sigma = 1.5$.



- **Question:** 0.20 is P($X >$?)
- **Response:**

Strategies for Normal Probability Problems

- Estimate probability given non-standard x
  - Standardize to z
  - Estimate probability using Rule
- Estimate non-standard x given probability
  - Estimate z
  - Unstandardize to x

Lecture Summary
(Normal Random Variables)

- Relevance of normal distribution
- Continuous random variables; density curves
- 68-95-99.7 Rule for normal R.V.s
- Standardizing/unstandardizing
- Probability problems
  - Find probability given z
  - Find z given probability
  - Find probability given x
  - Find x given probability