Chapter 6
Normal Probability Distributions

The Standard Normal Distribution

Applications of Normal Distributions

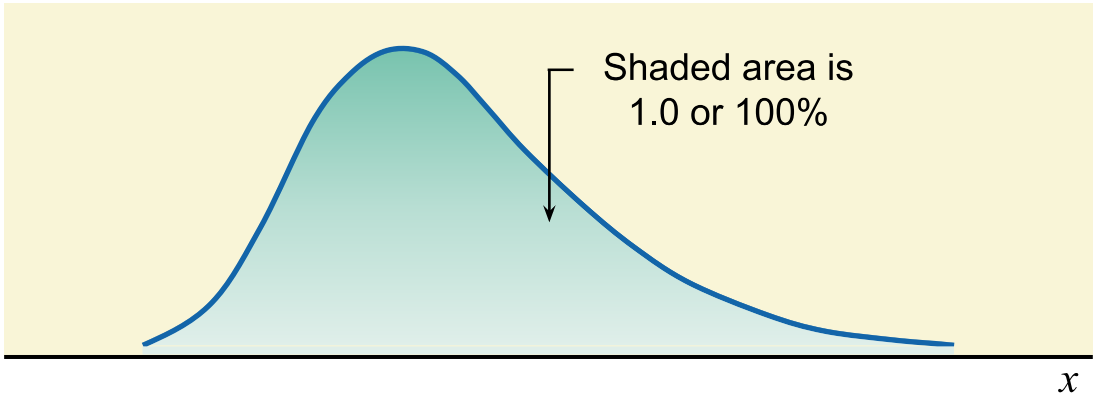
The Central Limit Theorem

Normal as Approximation to Binomial

Density Curve

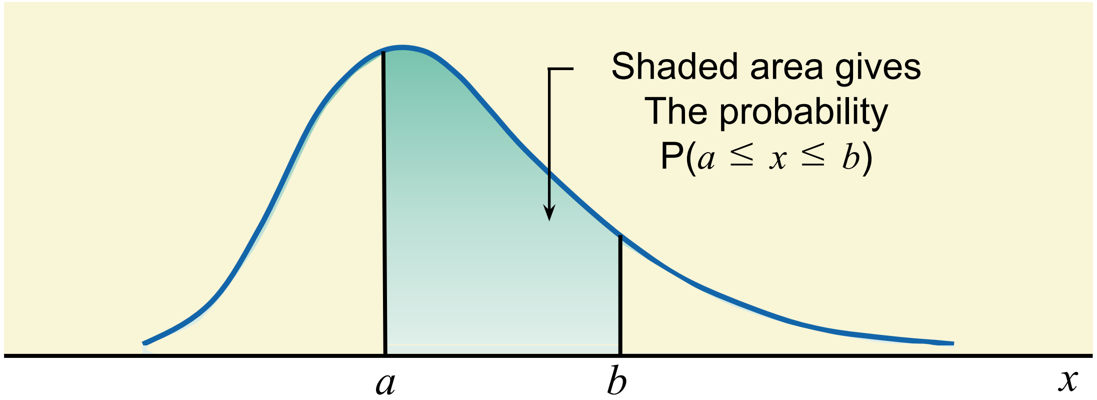
A density curve is the graph of a probability distribution of a continuous random variable. It must satisfy the following properties:

1. The total area under the curve = 1.
2. Every point on the curve must have a vertical height that is 0 or greater.



Area and Probability

Because the total area under the density curve is equal to 1, there is a correspondence between ***area*** and ***probability***.



$$P(a \leq X \leq b) = \text{Area under the curve from } a \text{ to } b$$

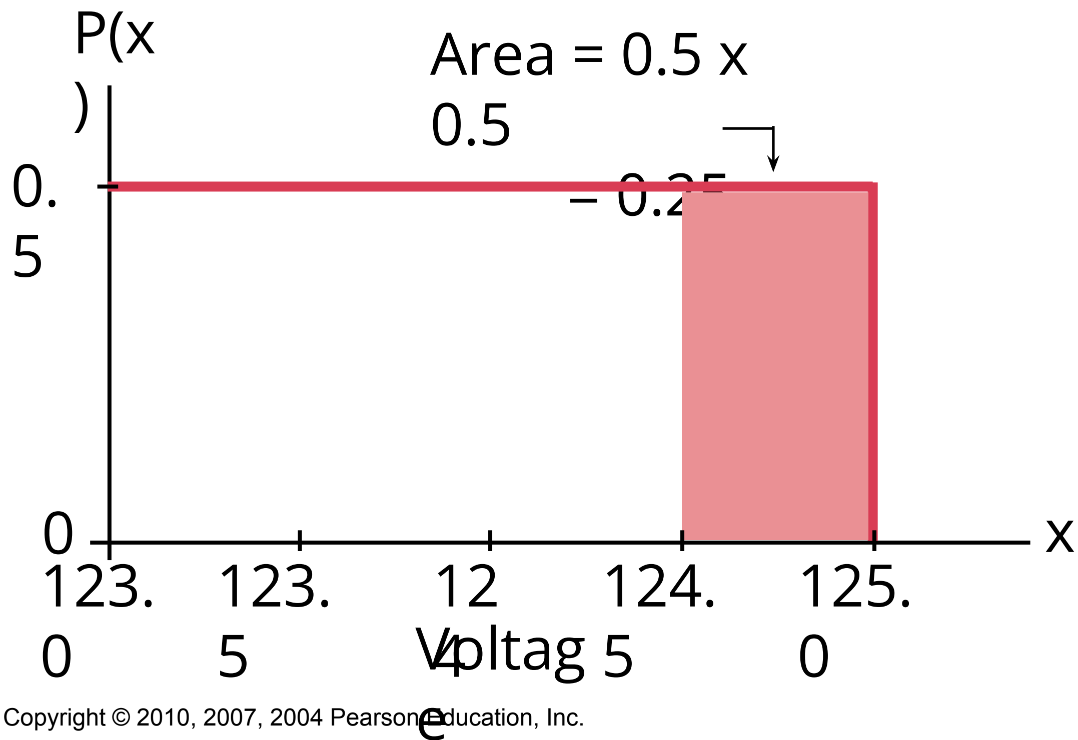
Uniform Distribution

(Definition) A continuous random variable has a uniform distribution if its values are spread evenly over the given range of an interval.

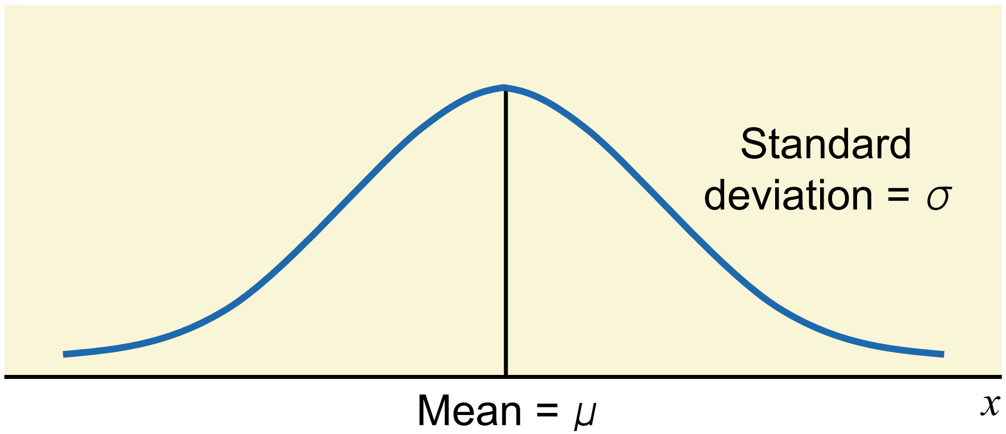
Note : the graph of a uniform distribution results in a rectangular shape.

(Example) A power company provides electricity with voltage levels between 123.0 volts and 125.0 volts, and all of the possible values are equally likely. Then, the voltage levels are uniformly distributed between 123.0 volts and 125.0 volts.

Find the probability that a randomly selected voltage level is greater than 124.5 volts.



Normal Probability Distribution



Density function

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} \right)$$

$$X \sim N(\mu, \sigma^2)$$

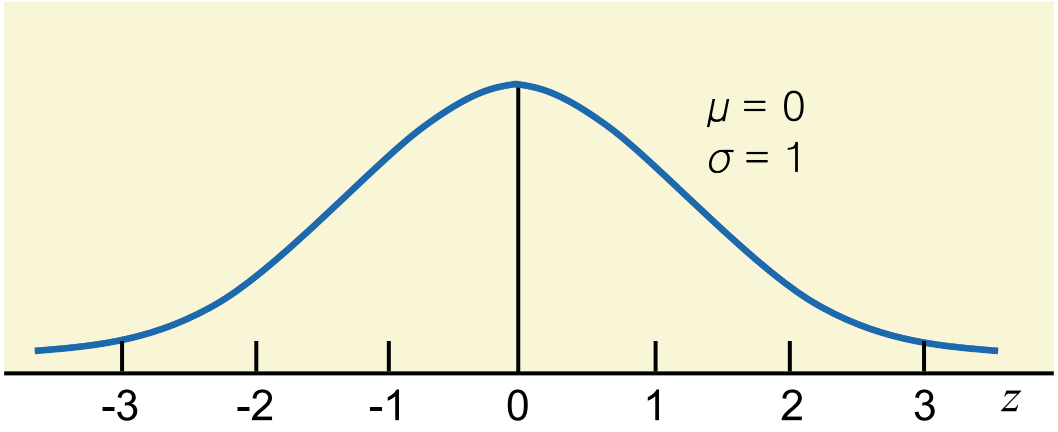
Two parameters

Properties

1. A bell–shaped curve
2. The total area under the curve is 1.0.
3. The curve is symmetric about the mean.

Standard Normal Distribution

The standard normal distribution is a normal probability distribution with $\underline{\mu = 0}$ $\underline{\text{and } \sigma = 1}$.



z–scores : The values marked on the horizontal axis of the standard normal curve

Finding probability when z-scores are given

P($a < Z < b$) = **probability that the *z*-score is between *a* and *b*.**

P($Z > a$) = **probability that the *z*-score is greater than *a*.**

P($Z < a$) = **probability that the *z*-score is less than *a*.**

To find the above probabilities,

we can use R, Excel or Statistical Table

R for normal distribution

- dnorm(x, mean = 0, sd = 1) =density function, not P(X=0)
- **pnorm**(x, mean = 0, sd = 1)=P(X<=x) :
- **qnorm**(p, mean = 0, sd = 1) : inverse function
- **Arguments**
  - **x** vector of quantiles.
  - **p** vector of probabilities.
  - **mean** vector of means.
  - **sd** vector of standard deviations.

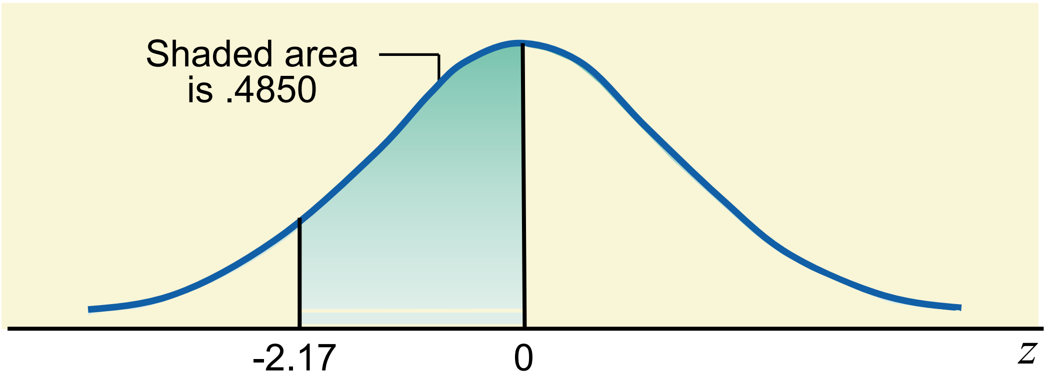
Example

z	.00	.01	...	$\boxed{.05}$	...	.09
-3.4	.0003	.0003	...	.0003	...	.0002
-3.3	.0005	.0005	...	.0004	...	.0003
-3.2	.0007	.0007	...	.0006	...	.0005
.	.	.	...	.	...	.
.	.	.	...	.	...	.
.	.	.	...	.	...	.
$\boxed{1.9}$	.9713	.9719	...	$\boxed{.9744}$	...	.9767
.	.	.	...	.	...	.
.	.	.	...	.	...	.
.	.	.	...	.	...	.
3.4	.9997	.9997	...	.9997	...	.9998

Required area

$$P(z < 1.95) = 0.9744$$

Example



$$P(-2.17 \le z \le 0)$$
$$= P(z \le 0) - P(z \le -2.17)$$
$$= 0.5 - 0.0150 = 0.4850$$

Table 6.3 **Area Under the Standard Normal Curve**

z	$\boxed{00}$	.01	…	$\boxed{.07}$	…	.09
-3.4	.0003	.0003	…	.0003	…	.0002
-3.3	.0005	.0005	…	.0004	…	.0003
-3.2	.0007	.0007	…	.0006	…	.0005
.	.	.	…	.	…	.
.	.	.	…	.	…	.
$\boxed{-2.1}$	.0179	.0174	…	$\boxed{.0150}$	…	.0143
.	.9713	.9719	…	.9744	…	.
.	.	.	…	.	…	.
$\boxed{0.0}$	$\boxed{.5000}$	.5040	…	.5279	…	.5359
.	.	.	…	.	…	.
.	.	.	…	.	…	.
3.4	.9997	.9997	…	.9997	…	.9998

Area to the left of z = 0

Area to the left of z = -2.17

Example I

Assume that the readings of a thermometer are normally distributed with the mean 0ºC and the standard deviation 1.00ºC. If one thermometer is randomly selected, find the probability that, at the freezing point of water (0º), the reading is less than 1.27º.

pnorm(1.27,0,1)

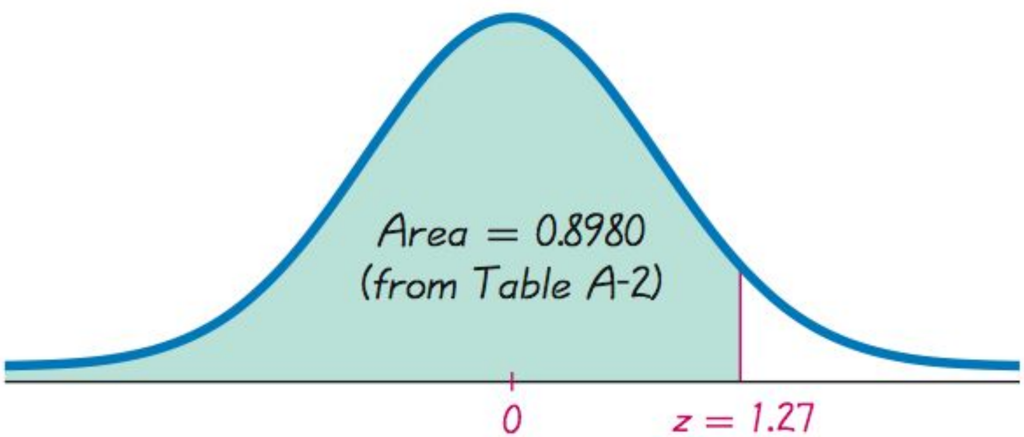
Example I

P(Z < 1.27) = ??

TABLE A-2	(Continued) Cumulative Area from the LEFT							
z	.00	.01	.02	.03	.04	.05	.06	.07
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675
0.2	.5793	.5382	.5871	.5910	.5948	.5987	.6026	.6064
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292

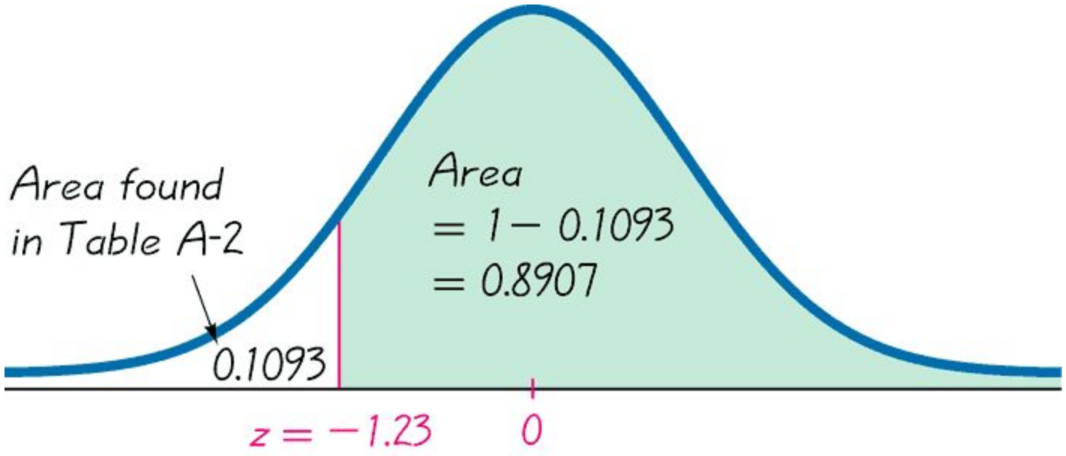
P (Z < 1.27) = 0.8980

The *probability* of randomly selecting a thermometer with a reading less than 1.27º is 0.8980.



Example II

If one thermometer is randomly selected, find the probability that it reads, at the freezing point of water, above –1.23 degrees.

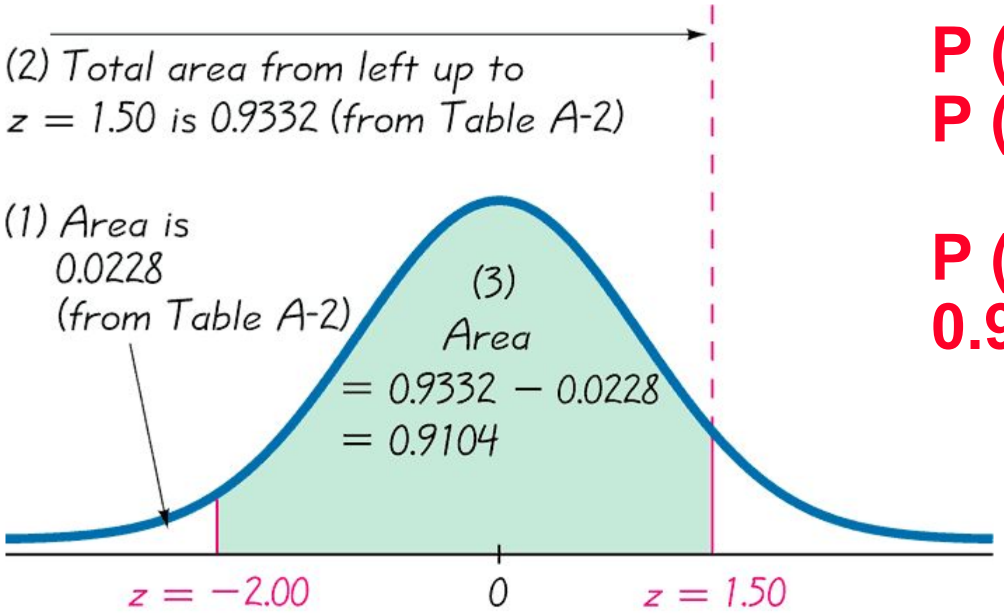


P (Z > –1.23)
=1-P(Z<=-1.23)
= 0.8907

1-pnorm(-1.23,0,1)

Example III

A thermometer is randomly selected. Find the probability that it reads (at the freezing point of water) between –2.00 and 1.50 degrees.

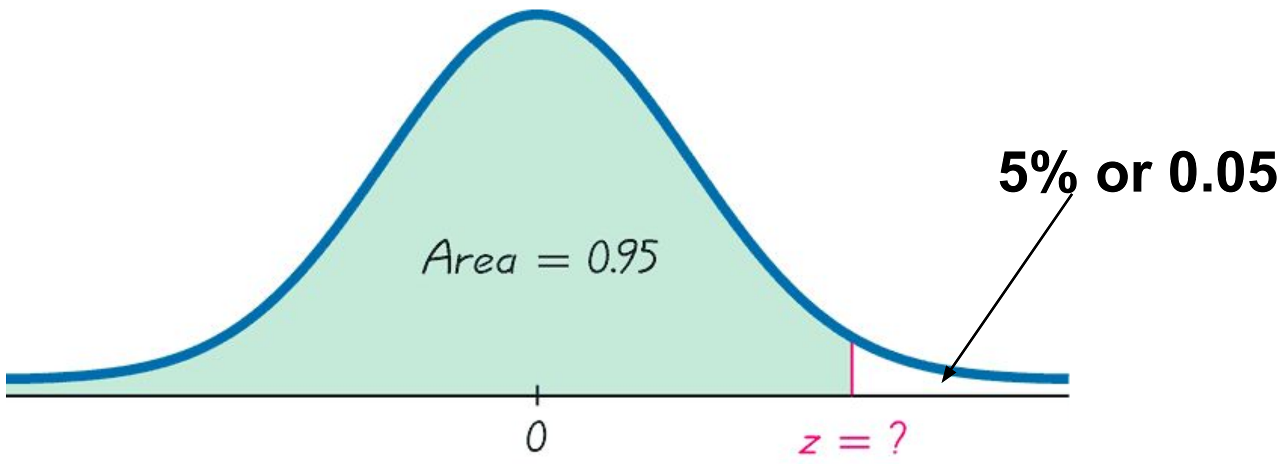


$P(Z < -2.00) = 0.0228$
$P(Z < 1.50) = 0.9332$

$P(-2.00 < z < 1.50) =$
$0.9332 - 0.0228 = 0.9104$

pnorm(1.5,0,1)
-pnorm(-2.00,0,1)

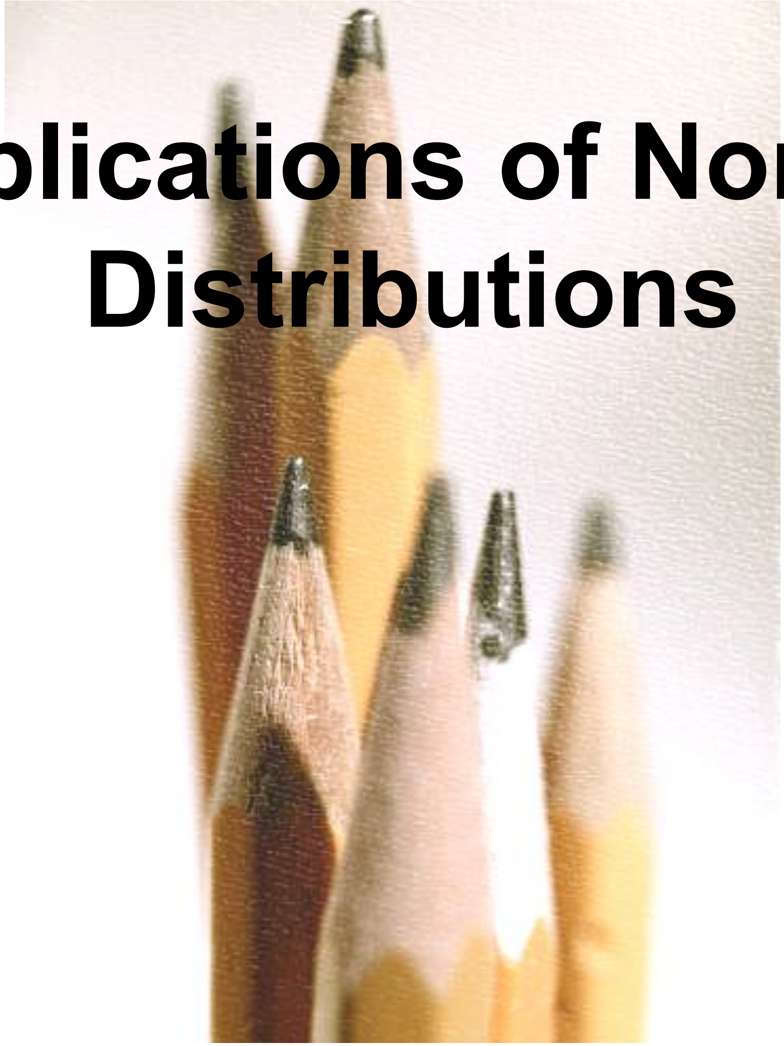
Finding z Scores When Given Probabilities – Inverse problem



Finding the 95th Percentile

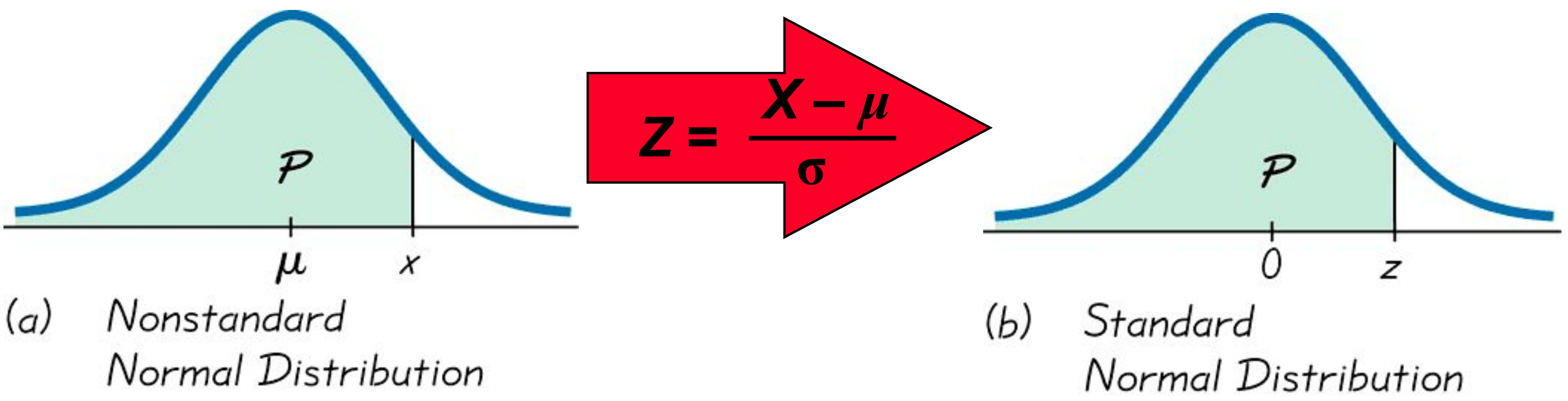
qnorm(0.95, mean = 0, sd = 1)=1.645

Applications of Normal Distributions



Converting to a Standard Normal Distribution

Conversion Formula : $$z = \frac{x - \mu}{\sigma}$$



$$\text{Suppose } X \sim N(\mu, \sigma^2), \quad \text{then } Z = \frac{X - \mu}{\sigma} \sim N(0,1)$$

$$E(Z) = \frac{1}{\sigma} E(X) - \frac{\mu}{\sigma} = 0$$

$$V(Z) = \frac{1}{\sigma^2} V(X) = 1$$

Example

Suppose that the weights of the men are normally distributed with a mean of 172 pounds and standard deviation of 29 pounds. If one man is randomly selected,
what is the probability he weighs less than 174 pounds?

(Solution)

Suppose $\boldsymbol{X} \sim \boldsymbol{N}(\boldsymbol{\mu}, \sigma^2)$, $\boldsymbol{\mu} = \mathbf{172}$, $\sigma = \mathbf{29}$.

Find P ($\boldsymbol{X}$ < 174).

Use

$$Z = \frac{X - \mu}{\sigma}$$

Example

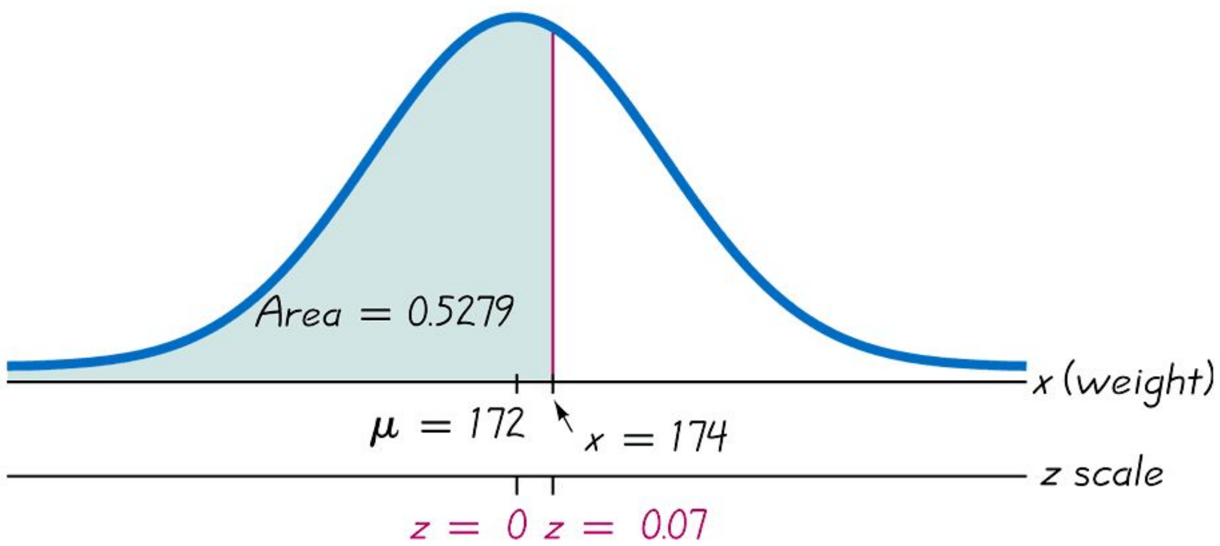
$\mu = 172$

$\sigma = 29$

$$Z = \frac{X - \mu}{\sigma}$$

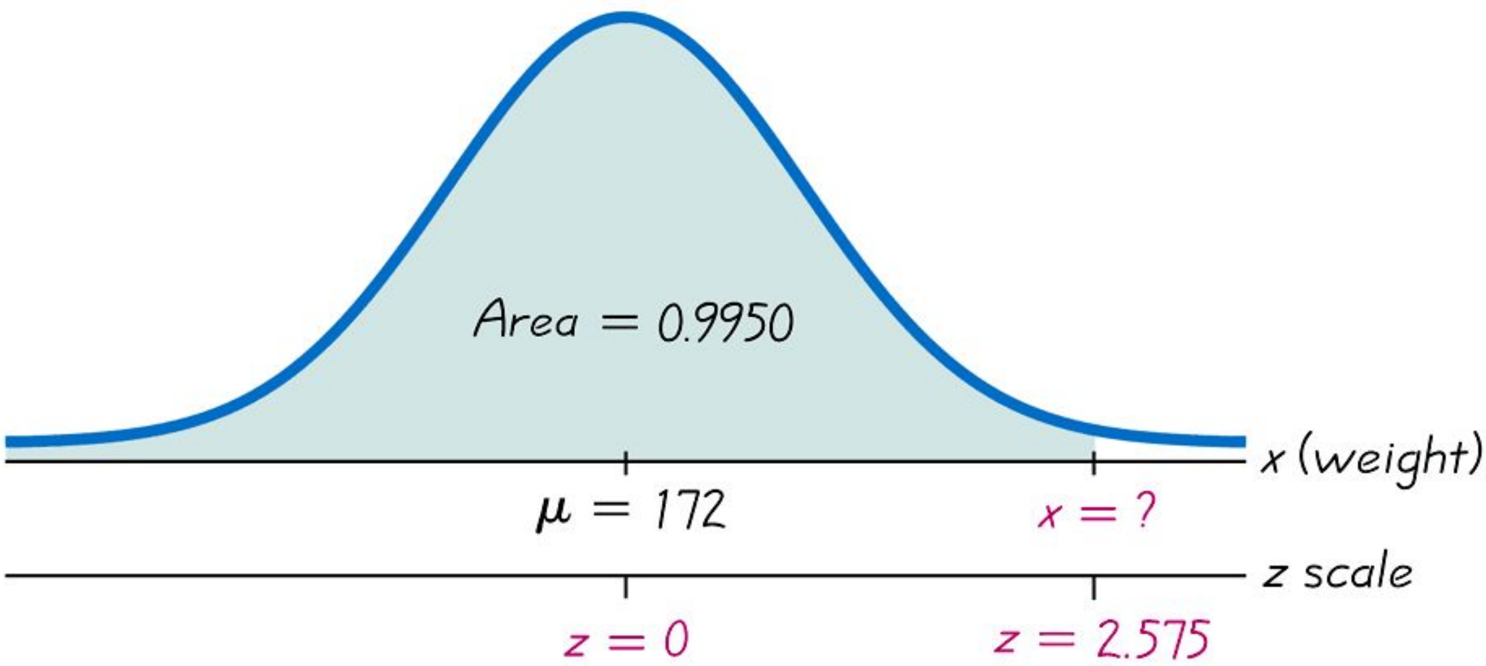
$$Z = \frac{174 - 172}{29} = 0.07$$

P (*X* < 174) = P(Z < 0.07)
= 0.5279



Example – inverse problem

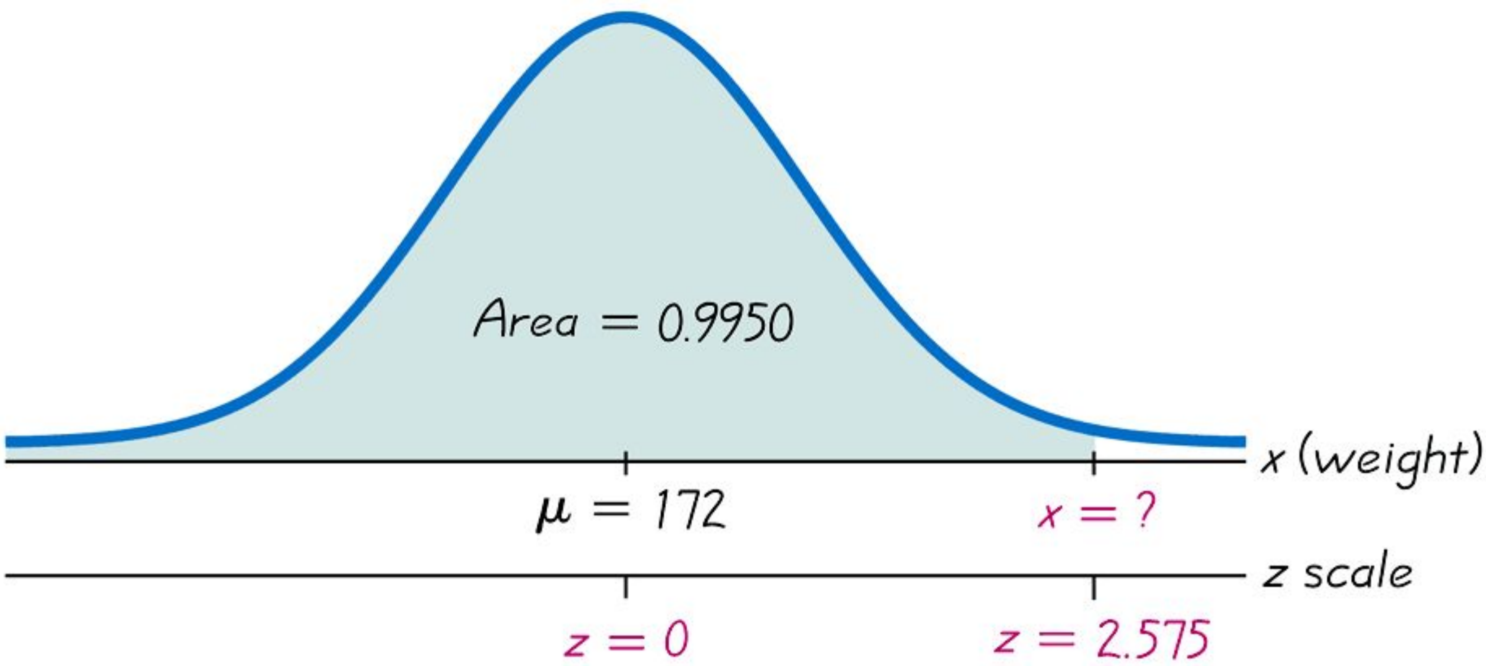
Use the data from the previous example to determine <u>what weight</u> separates the lightest 99.5% from the heaviest 0.5%?



Example – inverse problem

$x = \mu + (z * \sigma)$
$= 172 + (2.575 * 29)$
$= 246.675$

x=qnorm(0.995,172,29)
Or x= $\mu + (z * \sigma)$ *where*
z=qnorm(0.995,0,1)

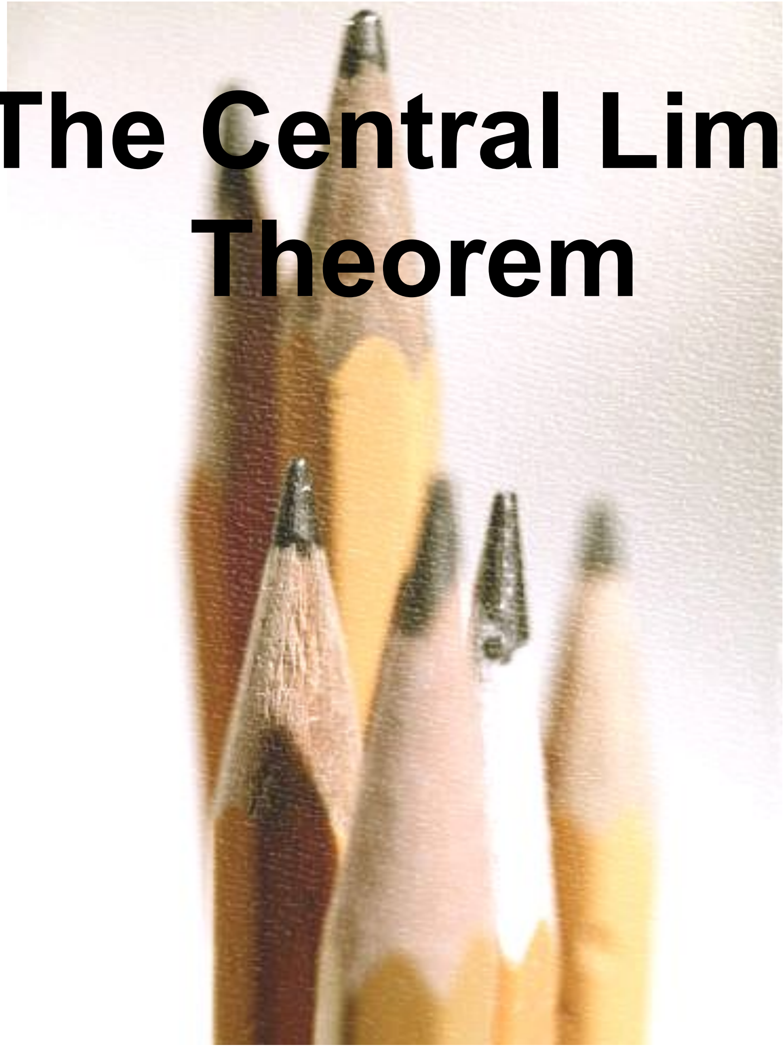


Sum of Independent Normal Random Variables

- Let X_1 and X_2 are <u>independent</u> and <u>normally distributed</u> with means μ_1 and μ_2, and variances σ_1^2 and σ_2^2, respectively. Then **their sum** $X = X_1 + X_2$ Is also **normally distributed** with mean $\mu_1 + \mu_2$ and variance $\sigma_1^2 + \sigma_2^2$

Prove this !

The Central Limit Theorem

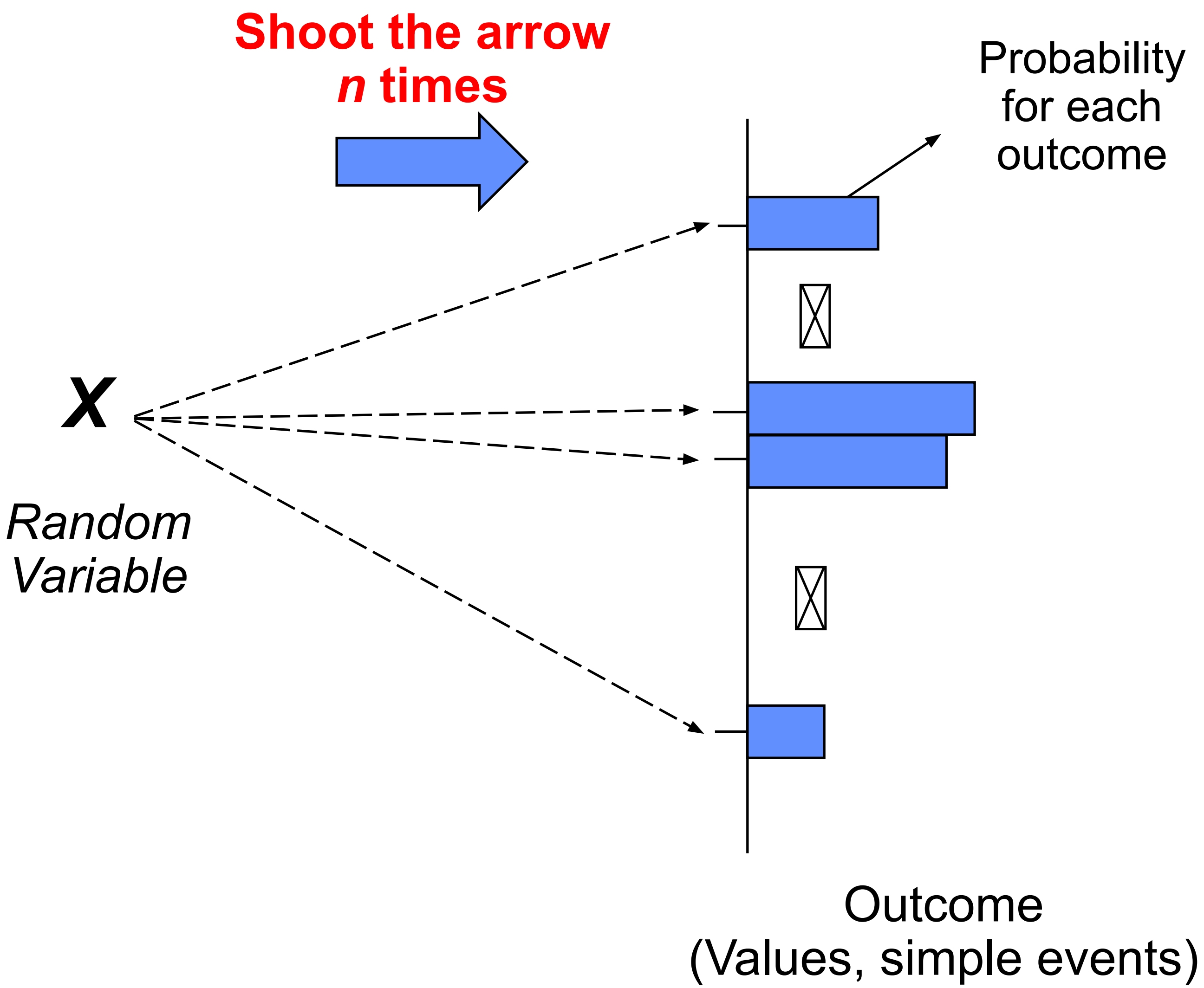


Key Concept

The *Central Limit Theorem* tells us that for a population with *any* distribution, the distribution of the sample means approaches a normal distribution as the sample size increases.

The procedure in this section form the foundation for estimating population parameters and hypothesis testing.

Random sample of size *n* from probability distribution



Central Limit Theorem

Given:

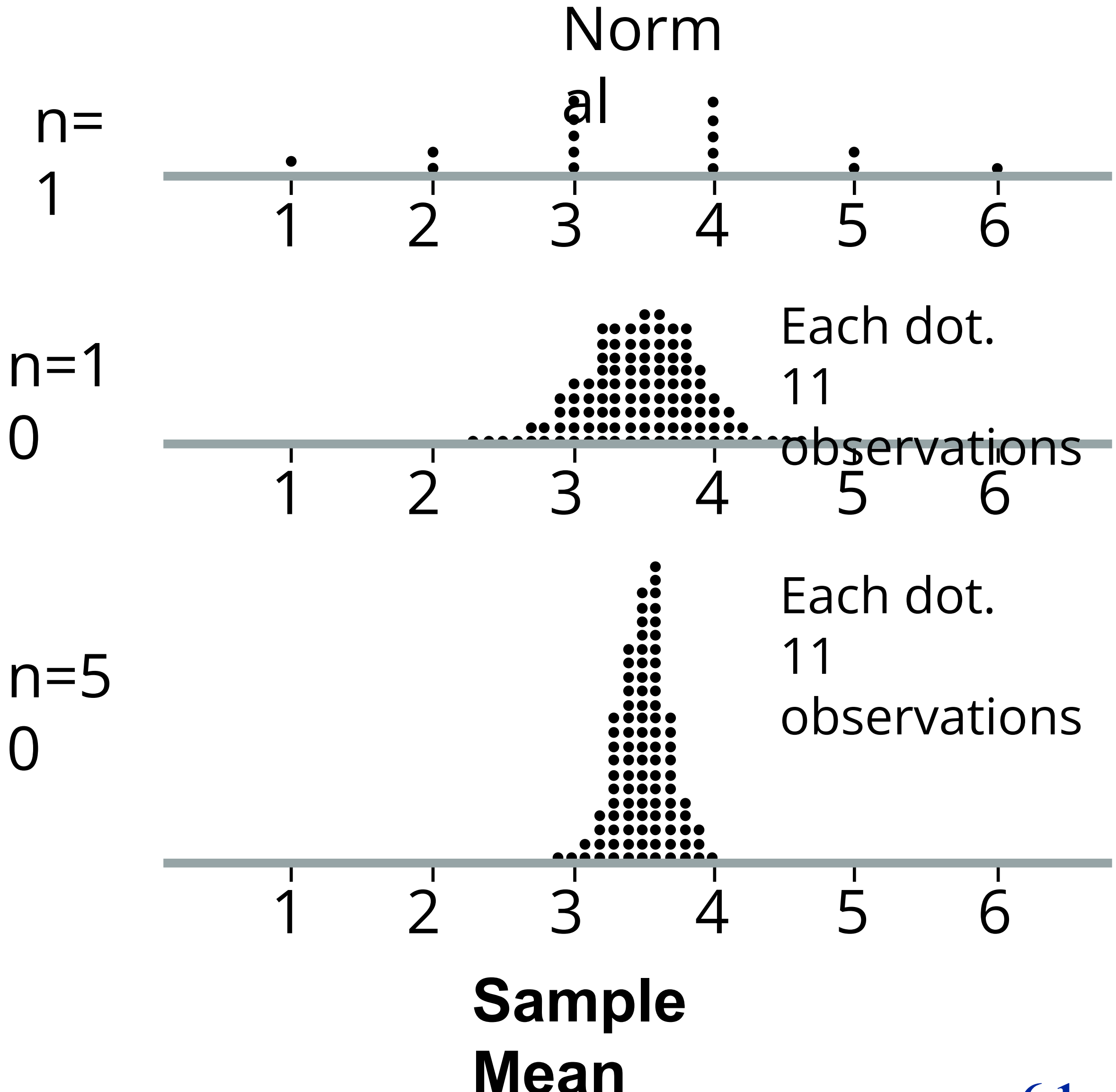
1. The random variable X has a distribution with mean μ and standard deviation σ (not necessarily be normal)
2. Simple random samples (all of size n) are selected from the population.

Conclusions:

1. The distribution of *sample means* will approach a normal distribution as the sample size increases,
2. The mean of the sample means = population mean μ.
3. The standard deviation of the sample means = $\sigma/\sqrt{n}$.

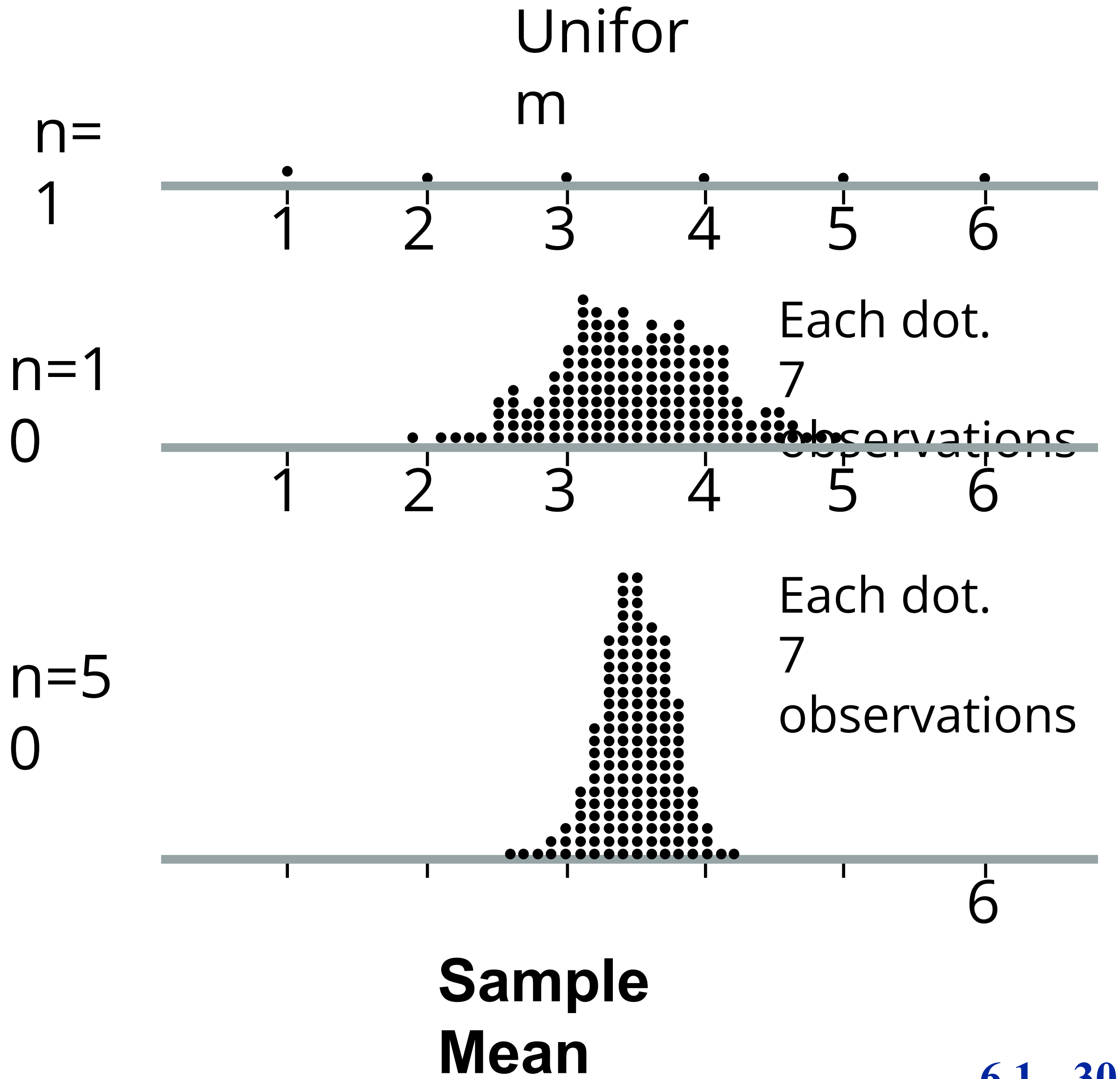
Example - Normal Distribution

As we proceed from *n* = 1 to *n* = 50, we see that the <u>distribution of sample means</u> is approaching the shape of a normal distribution.



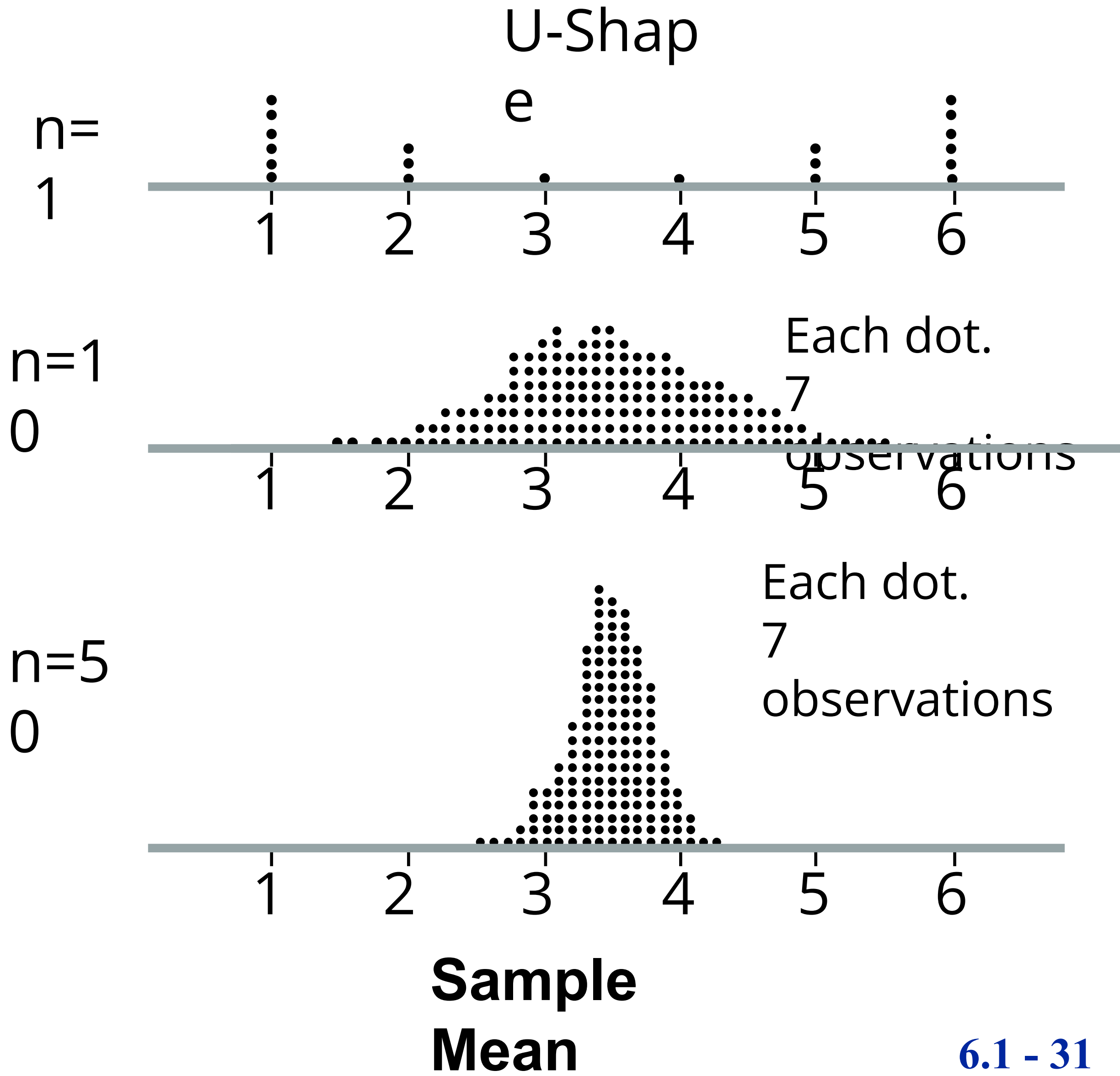
Example - Uniform Distribution

As we proceed from *n* = 1 to *n* = 50, we see that the <u>distribution of sample means</u> is approaching the shape of a normal distribution.



Example - U-Shaped Distribution

As we proceed from *n* = 1 to *n* = 50, we see that the <u>distribution of sample means</u> is approaching the shape of a normal distribution.



Notation

the mean of the <u>sample mean</u>

$$\mu_{\overline{X}} = E(\overline{X}) = \mu$$

the standard deviation of <u>sample mean</u>

$$\sigma_{\overline{X}} = \sigma(\overline{X}) = \frac{\sigma}{\sqrt{n}}$$

Show them !

Central Limit Theorem

Informal: Whatever the population, the distribution of $\bar{X}$ is normal with mean μ and standard deviation $\sigma/\sqrt{n}$ when *n* is large.

Formal: Whatever the population, approximately

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0,1)$$

if *n* is large (typically $n \geq 30$).

Practical Rules Commonly Used

(Case 1) The original population is *normally distributed*.

For any sample size n, the sample means will be normally distributed.

(Case 2) The original population is not *normally distributed*.

If samples of size $n \geq 30$, the distribution of the sample means can be approximated well by a normal distribution. The approximation gets closer to a normal distribution as the sample size n becomes larger.

If samples of size $n < 30$, central limit theorem may not be applied.

Example

Assume the population of weights of men is <u>normally distributed</u> with a mean of 172 lb and a standard deviation of 29 lb.

a) **Find the probability that if an *individual* man is randomly selected, his weight is greater than 175 lb.**

b) **Find the probability that *20 randomly selected men* will have a mean weight that is greater than 175 lb (so that their total weight exceeds the safe capacity of 3500 pounds).**

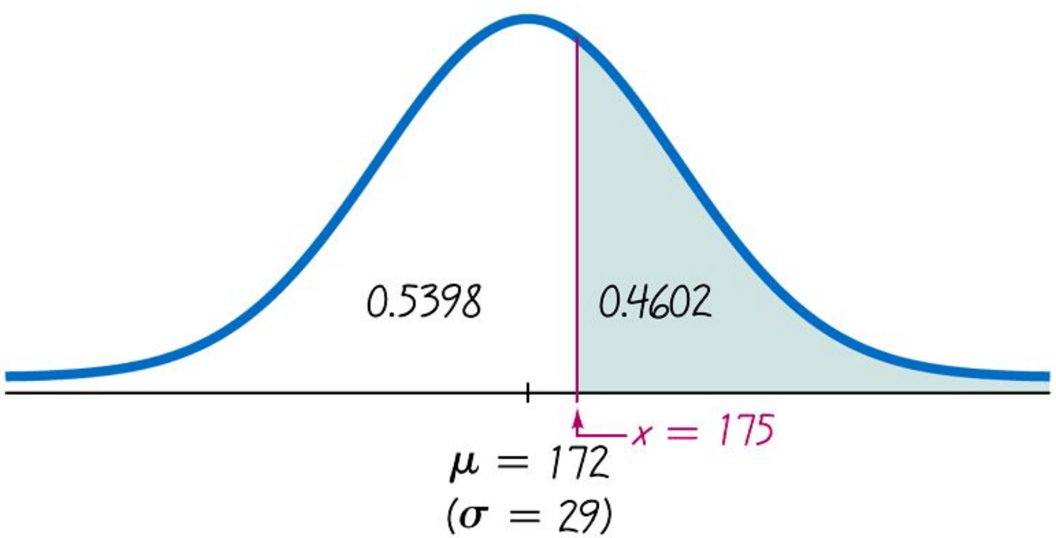
Example

a) Find the probability that if an *individual* man is randomly selected, his weight is greater than 175 lb.

$z = \dfrac{175 - 172}{29} = 0.10$

$P(z \leq 0.10) = 0.5398$

$P(x > 175) = 1 - P(z \leq 0.10) = 0.4602$



(a)

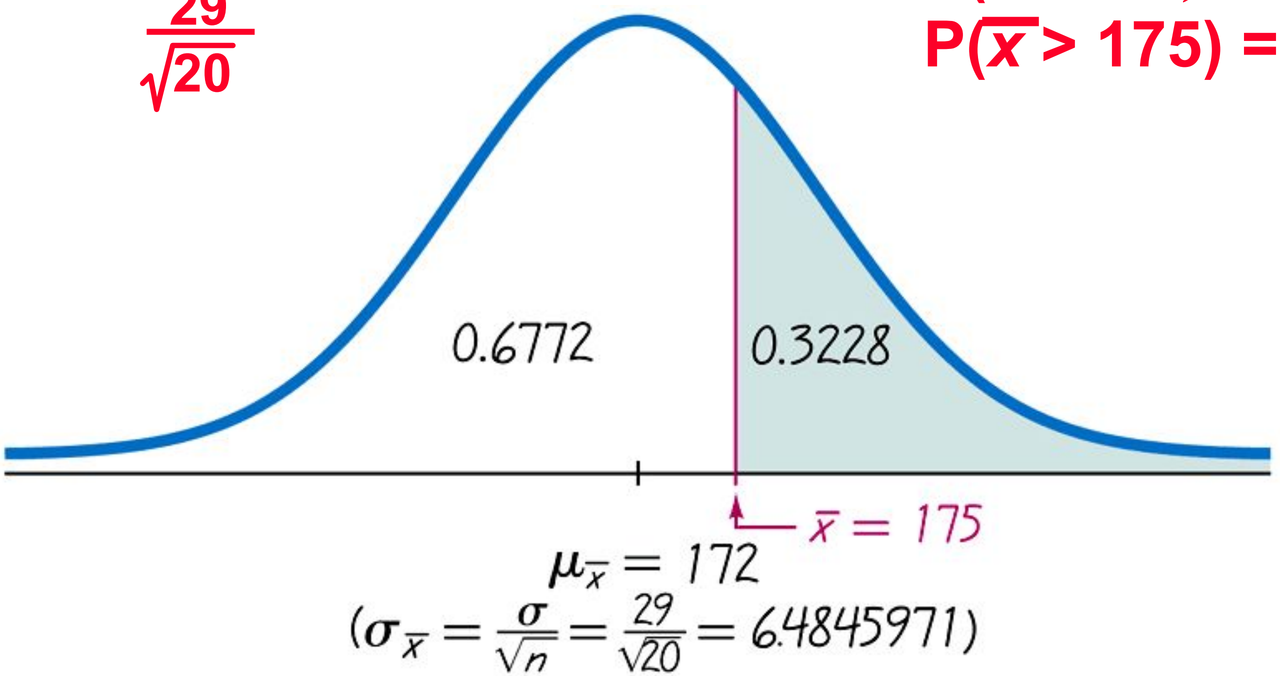
Example

b) Find the probability that *20 randomly selected men* will have a mean weight that is greater than 175 lb.

$$z = \frac{175 - 172}{\frac{29}{\sqrt{20}}} = 0.46$$

$P(z \leq 0.46) = 0.6772$

$P(\overline{x} > 175) = 0.3228$



(b)

Co

Example

Assume the population of weights of men has a mean of 172 lb and a standard deviation of 29 lb (not necessarily be *normal*). Find the probability that 30 randomly selected men will have a mean weight that is greater than 175 lb.

Normal as Approximation to Binomial

Review

Binomial Probability Distribution

1. The procedure must have a **fixed number of trials.**
2. The trials must be **independent.**
3. Each trial must have all outcomes classified into **two categories** (commonly, success and failure).
4. The probability of success remains the same in all trials.

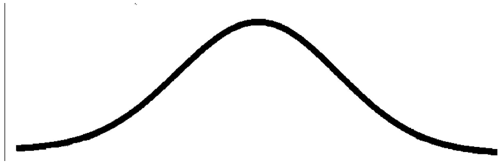
Approximation of a Binomial Distribution with a Normal Distribution

$np \geq 10$

$nq \geq 10$

then $\mu = np$ and $\sigma = \sqrt{npq}$

and the random variable has

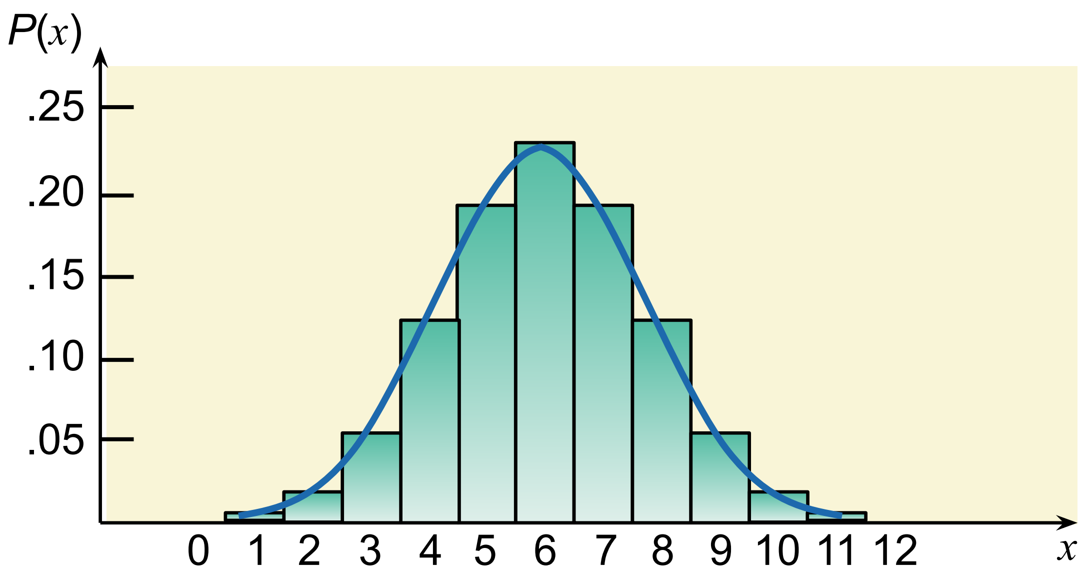
a  distribution.

(normal)

The Normal Approximation to the Binomial Distribution

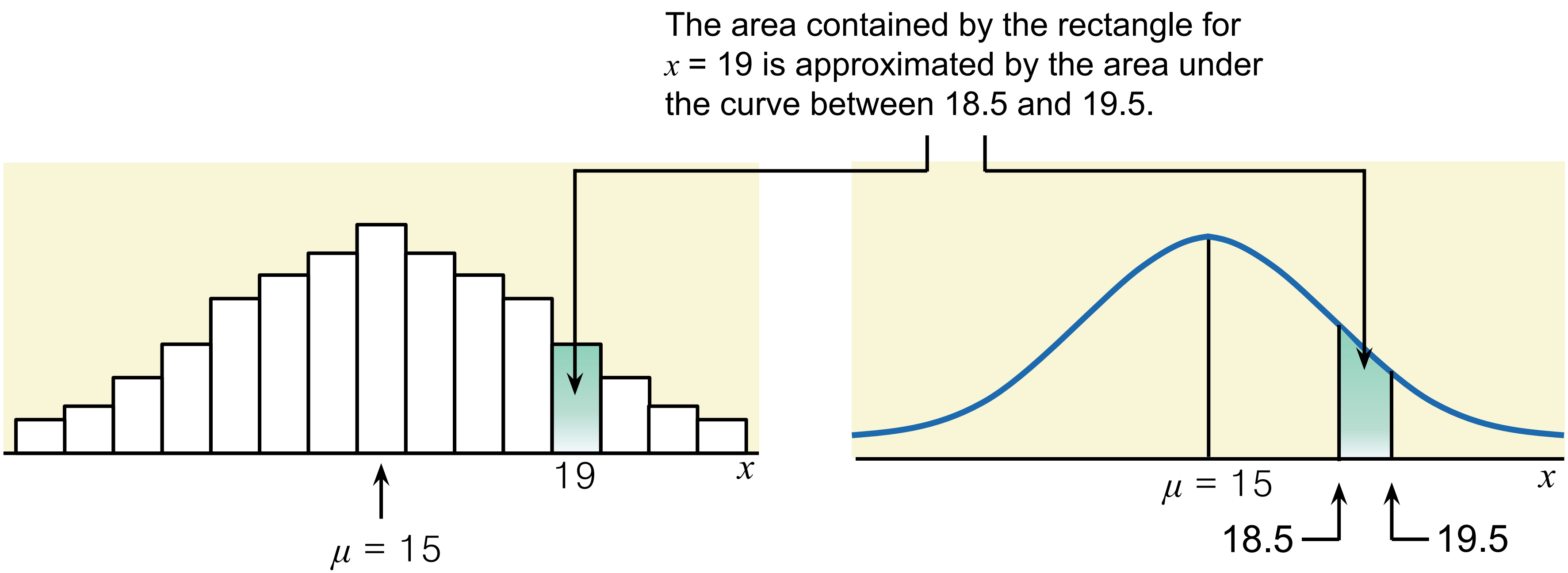
- Normal Distribution as an Approximation to Binomial Distribution

$$np > 10 \quad \text{and} \quad nq > 10$$



Why?

Continuity Correction



$$P(X = 19) \approx P(18.5 < Z < 19.5)$$

Procedure for Using a Normal Distribution to Approximate a Binomial Distribution

1. Check that $np \geq 10$ and $nq \geq 10$ before approximation
2. Calculate $\mu = np$ and $\sigma = \sqrt{npq}$.
3. Identify the discrete whole number x that is relevant to the binomial probability problem. Focus on this value temporarily.

Procedure for Using a Normal Distribution to Approximate a Binomial Distribution

4. **Draw a normal distribution centered about μ, then draw a *vertical strip area* centered over *x*. Mark the left side of the strip with the number equal to *x* – 0.5, and mark the right side with the number equal to *x* + 0.5. *Consider the entire area of the entire strip to represent the probability of the discrete whole number itself.***

5. **Determine whether the value of *x* itself is included in the probability. Determine whether you want the probability of at least *x*, at most *x*, more than *x*, fewer than *x*, or exactly *x*. Shade the area to the right or left of the strip; also shade the interior of the strip *if and only if x itself* is to be included. This total shaded region corresponds to the probability being sought.**

Example 1

Suppose there are 213 passengers in a train and the probability that a passenger is male is 0.5. Find the probability that there are "at least 122 men among 213 passengers by <u>using binomial distribution</u>

Example 1

Suppose there are 213 passengers in a train and the probability that a passenger is male is 0.5. Find the probability that there are "at least 122 men among 213 passengers by <u>using normal approximation</u>

$$\mu = np = 213 * 0.5 = 106.5$$

$$\sigma = \sqrt{npq} = \sqrt{213 * 0.5 * 0.5} = 7.2973$$

$$P(X \geq 122) \approx P(Y > 121.5)$$

$$= P(Z > \frac{121.5 - 106.5}{7.2973}) = P(Z > 2.06)$$

Example 2

$$X \sim B(n, p)$$
$$p = 0.48, \quad n = 200$$
$$P(80 < X < 90)? \quad = P(81 \leq X \leq 89)$$

$$\mu = 200 * 0.48 = 96, \quad \sigma = \sqrt{200 * 0.48 * 0.52}$$
$$Y \sim N(96, 7.065^2)$$
$$\underline{P(80 < X < 90) = P(81 \leq X \leq 89) \approx P(80.5 < Y < 89.5)}$$