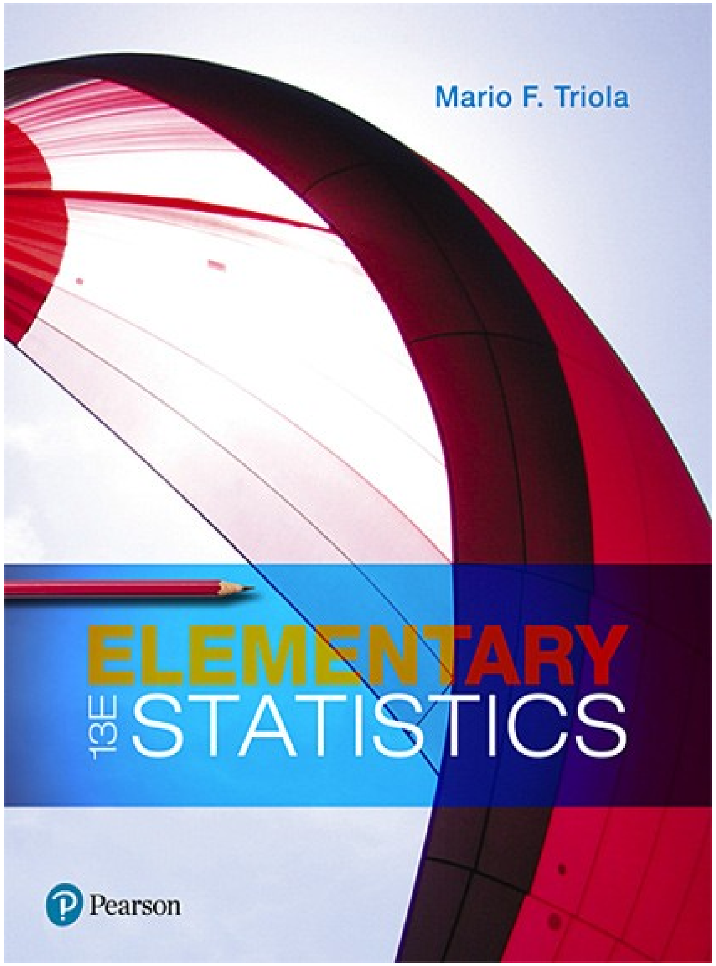
Elementary Statistics

Thirteenth Edition



Chapter 6

Normal Probability Distributions

Normal Probability Distributions

6-1 The Standard Normal Distribution

6-2 Real Applications of Normal Distributions

6-3 Sampling Distributions and Estimators

6-4 The Central Limit Theorem

6-5 Assessing Normality

6-6 Normal as Approximation to Binomial

Key Concept

In this section we introduce and apply the **central limit theorem.** The central limit theorem allows us to use a normal distribution for some very meaningful and important applications.

Central Limit Theorem

- Central Limit Theorem

  For all samples of the same size n with $n > 30$, the sampling distribution of $\bar{x}$ can be approximated by a normal distribution with mean μ and standard deviation $\dfrac{\sigma}{\sqrt{n}}$.

Central Limit Theorem and the Sampling Distribution *x*-bar

Given:

1. Population (with any distribution) has mean μ and standard deviation σ.
2. Simple random samples all of size n are selected from the population.

Practical Rules for Real Applications Involving a Sample Mean *x*-bar (1 of 2)

Requirements: Population has a normal distribution or *n* > 30:

Mean of all values of $\bar{x}$: $\mu_{\bar{x}} = \mu$

Standard deviation of all values of $\bar{x}$: $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

z score conversion of $\bar{x}$: $z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$

Practical Rules for Real Applications Involving a Sample Mean *x*-bar (2 of 2)

Original population is not normally distributed and $n \le 30$:

The distribution of $\bar{x}$ cannot be approximated well by a normal distribution, and the methods of this section do not apply.

Considerations for Practical Problem Solving (1 of 3)

1. **Check Requirements:** When working with the mean from a sample, verify that the normal distribution can be used by confirming that the original population has a normal distribution or the sample size is $n > 30$.

Considerations for Practical Problem Solving (2 of 3)

2. **Individual Value or Mean from a Sample?**

   Determine whether you are using a normal distribution with a **single** value x or the mean $\bar{x}$ from a sample of n values.

Individual value: When working with an **individual** value from a normally distributed population, use

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

Considerations for Practical Problem Solving (3 of 3)

2. **Individual Value or Mean from a Sample?**

   Determine whether you are using a normal distribution with a **single** value x or the mean $\bar{x}$ from a sample of n values.

   Mean from a sample of values: When working with a mean for some **sample** of n values, be sure to use the value of $\frac{\sigma}{\sqrt{n}}$ for the standard deviation of the sample means, so use

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}.$$

Notation for the Sampling Distribution of *x*-bar

If all possible simple random samples of size n are selected from a population with mean μ and standard deviation σ, the mean of all sample means is denoted by $\mu_{\bar{x}}$ and the standard deviation of all sample means is denoted by $\sigma_{\bar{x}}$.

Mean of all values of $\bar{x}$: $\mu_{\bar{x}} = \mu$

Standard deviation of all values of $\bar{x}$: $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$

Note: $\sigma_{\bar{x}}$ is called the **standard error of the mean** and is sometimes denoted as SEM.

Example: Safe Loading of Elevators (1 of 7)

The elevator in the car rental building at San Francisco International Airport has a placard stating that the maximum capacity is "4000 lb - 27 passengers." This converts to a mean passenger weight of 148 lb when the elevator is full. We will assume a worst-case scenario in which the elevator is filled with 27 adult males. Assume that adult males have weights that are normally distributed with a mean of 189 lb and a standard deviation of 39 lb.

a. Find the probability that 1 randomly selected adult male has a weight greater than 148 lb.

b. Find the probability that a sample of 27 randomly selected adult males has a mean weight greater than 148 lb.

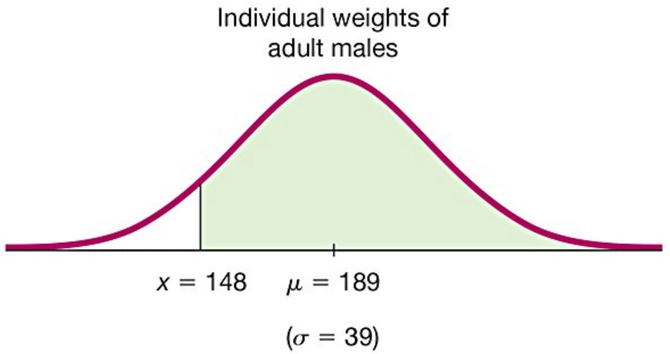
Example: Safe Loading of Elevators (2 of 7)

Solution (a)

Use the methods presented in a previous section because we are dealing with an **individual** value from a normally distributed population. We seek the area of the green-shaded region in the figure.

If using Table A-2, we convert the weight of x = 148 lb to the corresponding z score of z = −1.05, as shown here:

$$z = \frac{x - \mu}{\sigma} = \frac{148 - 189}{39} = -1.05$$

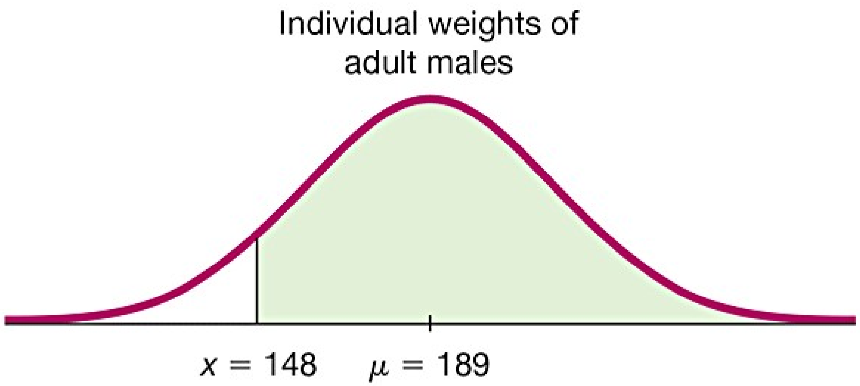


(a)

Example: Safe Loading of Elevators (3 of 7)

Solution (a)

We refer to Table A-2 to find that the cumulative area to the **left** of $z = -1.05$ is 0.1469, so the green-shaded area is $1 - 0.1469 = 0.8531$.



(a)

Pearson

Example: Safe Loading of Elevators (4 of 7)

Solution (b)

We can use the normal distribution if the original population is normally distributed or n > 30. The sample size is not greater than 30, but the original population of weights of males has a normal distribution, so samples of any size will yield means that are normally distributed.

Because we are now dealing with a distribution of sample means, we must use the parameters $\mu_{\bar{x}}$ and $\sigma_{\bar{x}}$, which are evaluated as follows:

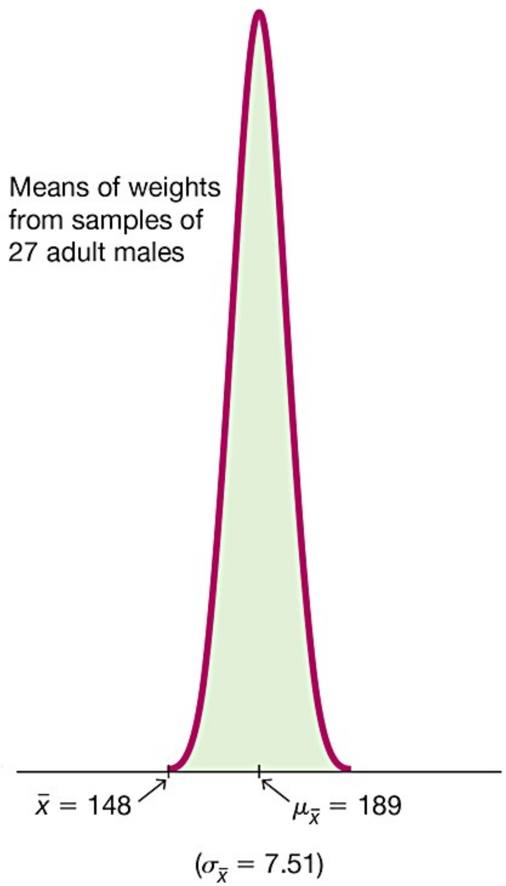
$$\mu_{\bar{x}} = \mu = 189$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{39}{\sqrt{27}} = 7.51$$

Example: Safe Loading of Elevators (5 of 7)

Solution (b)

We want to find the green-shaded area shown in the figure.



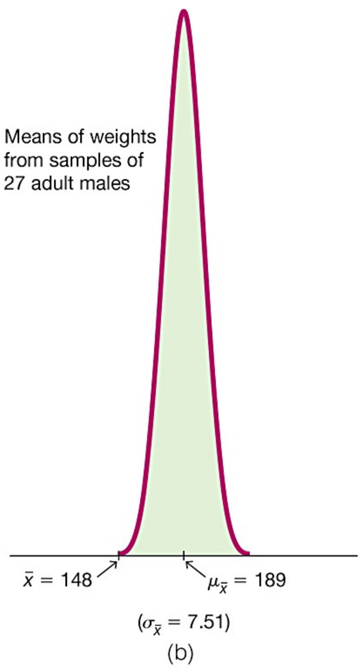
(b)

Example: Safe Loading of Elevators (6 of 7)

Solution (b)

If using Table A-2, we convert the value of $\bar{x}$ = 148 to the corresponding z score of z = −5.46, as shown here:

$$z = \frac{x - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{148 - 189}{\frac{39}{\sqrt{27}}} = \frac{-41}{7.51} = -5.46$$



From Table A-2 we find that the cumulative area to the left of z = −5.46 is 0.0001, so the green-shaded area is 1 − 0.0001 = 0.9999. We are quite sure that 27 randomly selected adult males have a mean weight greater than 148 lb.

Example: Safe Loading of Elevators (7 of 7)

Interpretation

There is a 0.8534 probability that an individual male will weigh more than 148 lb, and there is a 0.99999998 probability that 27 randomly selected males will have a mean weight of more than 148 lb.

Given that the safe capacity of the elevator is 4000 lb, it is almost certain that it will be overweight if it is filled with 27 randomly selected adult males.

Introduction to Hypothesis Testing

Identifying Significant Results with Probabilities: The Rare Event Rule for Inferential Statistics

If, under a given assumption, the probability of a particular observed event is very small and the observed event occurs **significantly less than or significantly greater than** what we typically expect with that assumption, we conclude that the assumption is probably not correct.

Example: Body Temperatures

Assume that the population of human body temperatures has a mean of 98.6°F, as is commonly believed. Also assume that the population standard deviation is 0.62°F (based on data from University of Maryland researchers). If a sample of size n = 106 is randomly selected, find the probability of getting a mean of 98.2°F or lower.

Example – Body Temperatures (1 of 4)

Solution

We work under the assumption that the population of human body temperatures has a mean of 98.6°F. We weren't given the distribution of the population, but because the sample size n = 106 exceeds 30, we use the central limit theorem and conclude that the distribution of sample means is a normal distribution with these parameters:

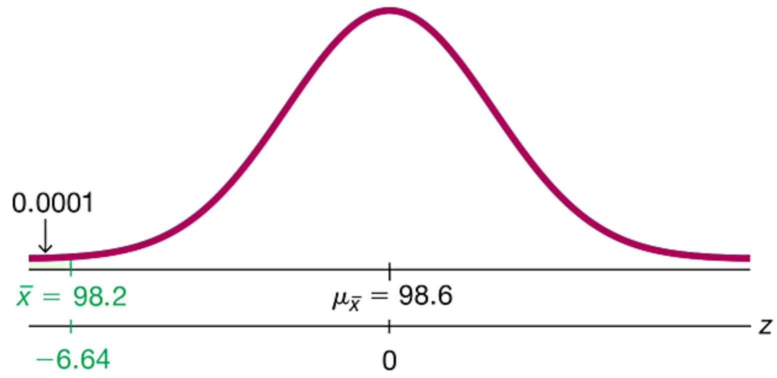
$$\mu_{\bar{x}} = \mu = 98.6 \text{ (by assumption)}$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.62}{\sqrt{106}} = 0.0602197$$

Example – Body Temperatures (2 of 4)

Solution

The figure shows the shaded area (see the tiny left tail of the graph) corresponding to the probability we seek. Having already found the parameters that apply to the distribution shown in the figure, we can now find the shaded area by using the same procedures developed previously.

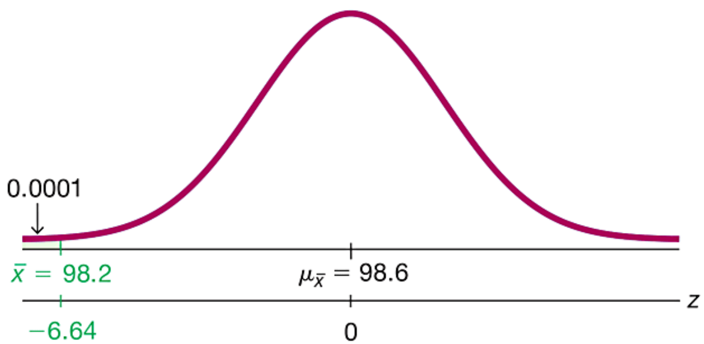


Example – Body Temperatures (3 of 4)

Solution

If we use Table A-2 to find the shaded area in the figure, we must first convert the score of x = 98.20°F to the corresponding z score:

$$z = \frac{x - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{98.20 - 98.6}{0.0602197} = -6.64$$



Referring to Table A-2, we find that $z = -6.64$ is off the chart, but for values of z below −3.49, we use an area of 0.0001 for the cumulative left area up to $z = -3.49$. We therefore conclude that the shaded region is 0.0001.

Example – Body Temperatures (4 of 4)

Interpretation

The result shows that if the mean of our body temperatures is really 98.6°F, as we assumed, then there is an extremely small probability of getting a sample mean of 98.2°F or lower when 106 subjects are randomly selected. University of Maryland researchers did obtain such a sample mean, and after confirming that the sample is sound, there are two feasible explanations: (1) The population mean really is 98.6°F and their sample represents a chance event that is extremely rare; (2) the population mean is actually lower than the assumed value of 98.6°F and so their sample is typical. Because the probability is so low, it is more reasonable to conclude that the population mean is lower than 98.6°F. In reality it appears that the true mean body temperature is closer to 98.2°F!

Correction for a Finite Population

When sampling without replacement and the sample size n is greater than 5% of the finite population size N (that is, $n > 0.05N$), adjust the standard deviation of sample means $\sigma_{\overline{x}}$ by multiplying it by this **finite population correction factor:**

$$\sqrt{\frac{N-n}{N-1}}$$