vantes from *Don Quixote de la Mancha*

APTER

mples and Sampling stributions

Sampling Frame

A **sampling frame** is a list which identifies all members of the population.

DEFINITION

Simple Random Sample

A **simple random sample** from a finite population is one in which every possible sample of the same size n has the same probability of being selected.

DEFINITION

Biased Sample

A **sample is biased** if it overrepresents or underrepresents some segment(s) of the population.

DEFINITION

Other Forms of Sampling

Judgment Sample

A sample in which the observations are selected by an expert in the field and not picked at random.

DEFINITION

Convenience Sample

A sample of observations that are easily obtained and not random.

DEFINITION

Drawing samples that are representative of a population is often quite difficult. One "sampling method" that is often seen in the media is **voluntary sampling**. Usually some question is posed to a large audience and people "volunteer" to participate in the sample. The internet has made this kind of "sampling" popular. However, this sampling method can often produce misleading results.

Systematic Sample

A sample in which you choose a starting point and then every k^{th} member of the population is included in the sample.

DEFINITION

Cluster Sampling

Cluster sampling involves dividing the population into clusters, and randomly selecting a sample of clusters to represent the population. Cluster sampling is used when "natural" groupings are evident in the population.

DEFINITION

Bike Paths Survey To save time, the planner decided to use cluster sampling. The residential portion of the city was divided into 947 blocks, each containing 20 homes, as shown in Fig. 1.3. Explain how the planner used cluster sampling to obtain a sample of 300 homes.

Figure 1.3

Stratified Sampling

In **stratified sampling,** the population is divided into **strata**, which are subpopulations. A **strata** can be any identifiable characteristic that can be used to classify the population. If the population consists of people, then strata could be sex, income, political party, religion, education, race, or location.

DEFINITION

947

200	upper	21%	3
400	middle	42%	6
347	lower	37%	6

Sampling Distribution of a Statistic

The **sampling distribution of a statistic** (such as the sample mean or sample proportion) is the probability distribution of all values of the statistic when all possible samples of size *n* are taken from a population.

DEFINITION

Point Estimator

A **point estimator** is a single-valued estimate calculated from the sample data, which is intended to be close to the true population value.

DEFINITION

$\mu \to \bar{x}$

$\sigma^2 \to s^2$

$\sigma \to s$

Pop sample

Unbiased Estimators

1. The sample mean, $\bar{x}$, is an unbiased estimator of μ.
2. The sample proportion, $\hat{p}$, is an unbiased estimator of p.
3. The sample variance, s^2, is an unbiased estimator of σ^2.

PROPERTIES

The Central Limit Theorem

If a sufficiently large random sample (i.e., $n > 30$) is drawn from a population with mean μ and standard deviation σ, the distribution of the sample mean will have the following characteristics.

1. An approximately normal distribution regardless of the distribution of the underlying population.
2. $\mu_{\overline{x}} = E(\overline{x}) = \mu$ (The mean of the sample means equals the population mean.)
3. $\sigma_{\overline{x}} = \dfrac{\sigma}{\sqrt{n}}$ (The standard deviation of the sample means equals the standard deviation of the population divided by the square root of the sample size.)

DEFINITION

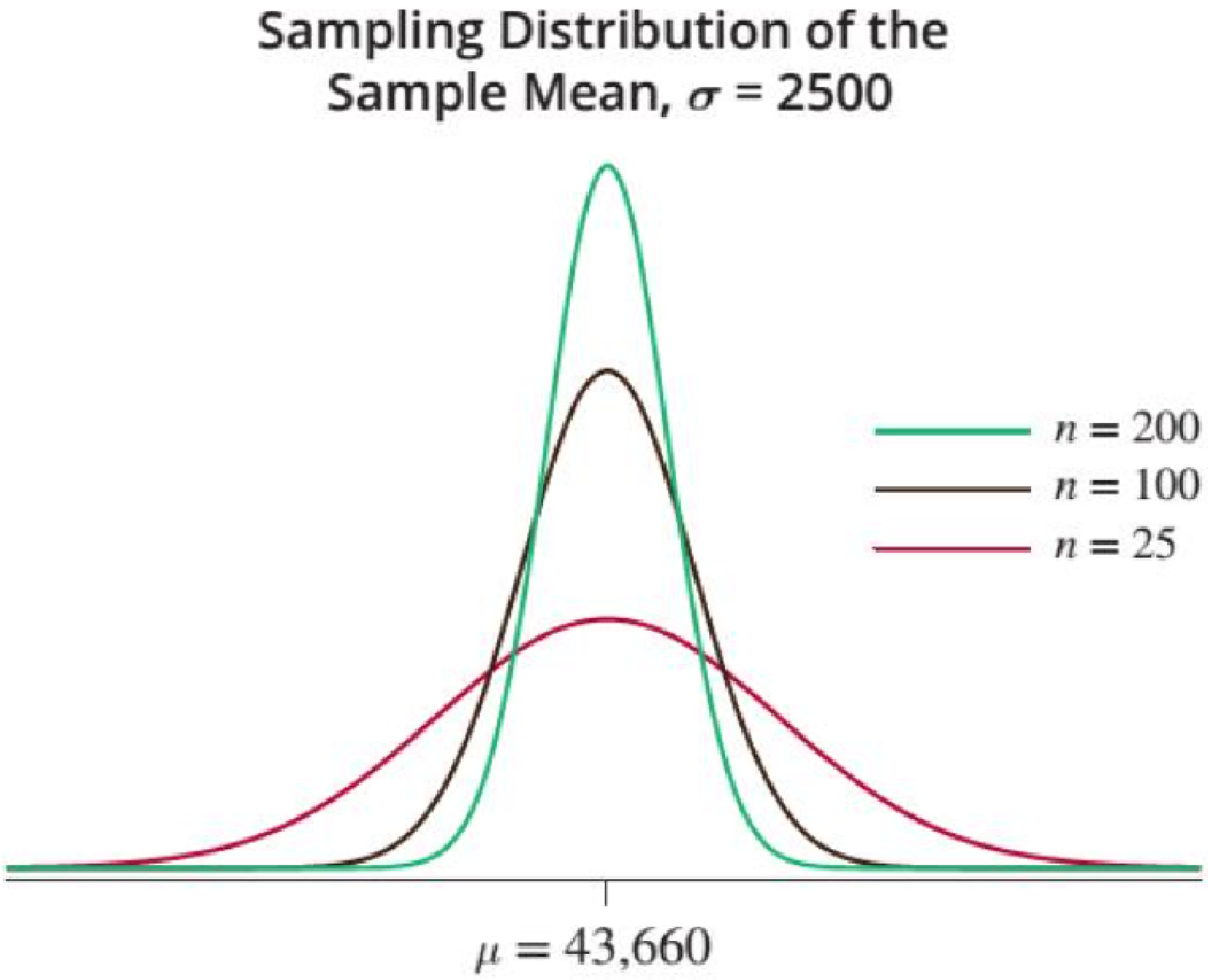
Sampling Distribution of the
Sample Mean, σ = 2500
$n = 200$
$n = 100$
$n = 25$
$\mu = 43,660$

Figure 9.3.2

A travel agency conducted a survey of the prices charged by ocean cruise ship lines and determined they were approximately normally distributed with a mean of $110 per day and a standard deviation of $20 per day.

a. If an ocean cruise ship line is chosen at random, find the probability that it will charge less than $99 per day.

b. What is the probability that the average charge for a randomly selected sample of 35 ocean cruise ship lines will be less than $99 per day?

$$z = \frac{99-110}{20} = \frac{-11}{20} = -0.55$$

$\rightarrow$ **0,29116**

b) $\sigma = \frac{20}{\sqrt{35}} = 3,38$ $\quad z = \frac{99-110}{3.38} = \frac{-11}{3.38} = -3,25$

0,00058

The average score for a water safety instructor (WSI) exam is 75 with a standard deviation of 12. Fifty scores for the WSI exam are randomly selected.

a. Find the probability that the average of the fifty scores is at least 80.

b. Find the probability that the average of the fifty scores is at most 70.

c. Find the probability that the average of the fifty scores is between 72 and 78.

$\mu = 75 = \mu_{\bar{x}}$

$\sigma = 12 \rightarrow \sigma_{\bar{x}} = \frac{12}{\sqrt{50}} = 1.70$

$n = 50$

$z = \frac{x - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$

a) $z > 2.94$

.00164

b) $z < -2.94$

.00164

c) $< z <$

A college food service buys frozen fish in boxes labeled 10 pounds. The true average weight of the boxes is 8 pounds with a standard deviation of 2 pounds. The food service director suspects that the boxes do not contain as much fish as advertised. He decides to inspect 40 boxes from the next shipment. If the average weight is less than 10 pounds he will reject the entire shipment. Find the probability that the food service director will not reject the shipment.

$\mu_{\bar{x}} = 8$ $\quad \sigma = 2$ $\quad \sigma_{\bar{x}} = \frac{2}{\sqrt{40}} = 0.32$

$z = \frac{10 - 8}{0.32} = \boxed{6.25}$

Sampling Distribution of the Sample Proportion

If the population is infinite and the sample is sufficiently large, the distribution of $\hat{p}$ has the following characteristics:

1. An approximately normal distribution.
2. $\mu_{\hat{p}} = E(\hat{p}) = p$. (The mean of the sample proportions equals the population proportion.)
3. $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} \approx \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$.

Suppose a sample of 400 persons is used to perform a taste test. If the true fraction in the population that prefers Pepsi is really 0.5, what is the probability that less than 0.44 of the persons in the sample will prefer Pepsi?

$$p = \mu = 0.5$$

$$\sigma = \sqrt{\frac{0.5(1-0.5)}{400}} = 0.025$$

$$z = \frac{0.44 - 0.5}{0.025} = -2.4 \rightarrow .00820$$

Suppose a sample of 500 is used to estimate the fraction of voters that favor a particular candidate. If the population proportion that favors the candidate is really 0.4, what is the probability that the error of estimation will be less than 0.05?

Approximately 7% of the nation's public-school children in grades 2 through 5 take medication for attention deficit hyperactivity disorder (ADHD), a developmental disorder characterized by impulsiveness or difficulty concentrating or sitting still. The main treatment prescribed for ADHD is Ritalin, a relatively safe drug with few side effects. A sample of 286 students is taken.

a. Find the probability that at least 4% of the school children in the sample take medication for ADHD.

b. Find the probability that between 5% and 8% of the school children in the sample take medication for ADHD.