

TA Office Hours, STATA class

- **TEACHING ASSISTANTS**

Tiffany Chou: tmchou@ucsd.edu

Michael Bauer: mbauer@ucsd.edu

Leah Nelson: lknelson@ucsd.edu

**OH: Monday 7-8PM, Center 119 ;
Wednesday 3:30-5, Seq. 244**

**** STATA Evening Class ****

**Jan 14: Monday 7-8PM or Monday 8-9PM;
(Econ 100 computer lab)**

Econometrics 120B

Lesson 2: Review of Probability

- 0. Last Class: 2 Themes**
- 1. Probability**
- 2. Random Variables**
- 3. Expected Value**
- 4. Variance, Covariance**
- 5. Random Sampling**
- 6. Sample size matters**
- 7. Working with Normal Distribution**
- 8. Central Limit Theorem**
- 9. Problem Set #1**

Next Class: Statistics Review

Review: Why Econometrics?

1. Why study Econometrics?

2. Who needs data anyway?

3. If you had some, what would you do with it?

Coffee example [Quantifying uncertainty]

GRE example [Correlation and causality]

These two themes will recur throughout the course

4. Types of data:

Experimental vs. nonexperimental data.

Cross-sections, Time-Series, Panels.

5. Syllabus & logistics - econ.ucsd.edu/~elib/120b

Next Class.. Statistics Review

[Don't forget: Problem Set #1](#)

1. Probability - events best thought of as uncertain

Events:

- A – Chargers beat Colts (0-no,1-yes)
- B - Chargers win the Super Bowl
- C - Iraq war ends by 2010
- D – Hilary Clinton wins Democratic nomination
- E - Angelina Jolie and Brad Pitt adopt a child in 2008



Joint, marginal and conditional probabilities.

	A	0	1	
B				
0		.3	.55	.85
1		0	.15	.15
		.3	.7	

- what's $P(A=1, B=1)$? 0.15
- $P(A=1)$?
- $P(B=1|A=1) = P(A=1, B=1)/P(A=1) = ?$ *.15 / .7 = .214*
- $P(A=1|B=1) = ?$
- $P(A=1|E=1) = ?$



*Independence of A and E
 $P(A=1|E=1) = P(A=1)$*

2. Random Variables

- **X is a random variable if for every real number a there exists a probability $P(X \leq a)$.**
eg's: X is touchdowns scored by LT in playoffs
q – coffee demanded
q(p) - coffee demanded at price p
- **Last class the predicted coffee demand, slope and intercept were random variables, why? because of sampling variation.**

Distribution (Density) Functions

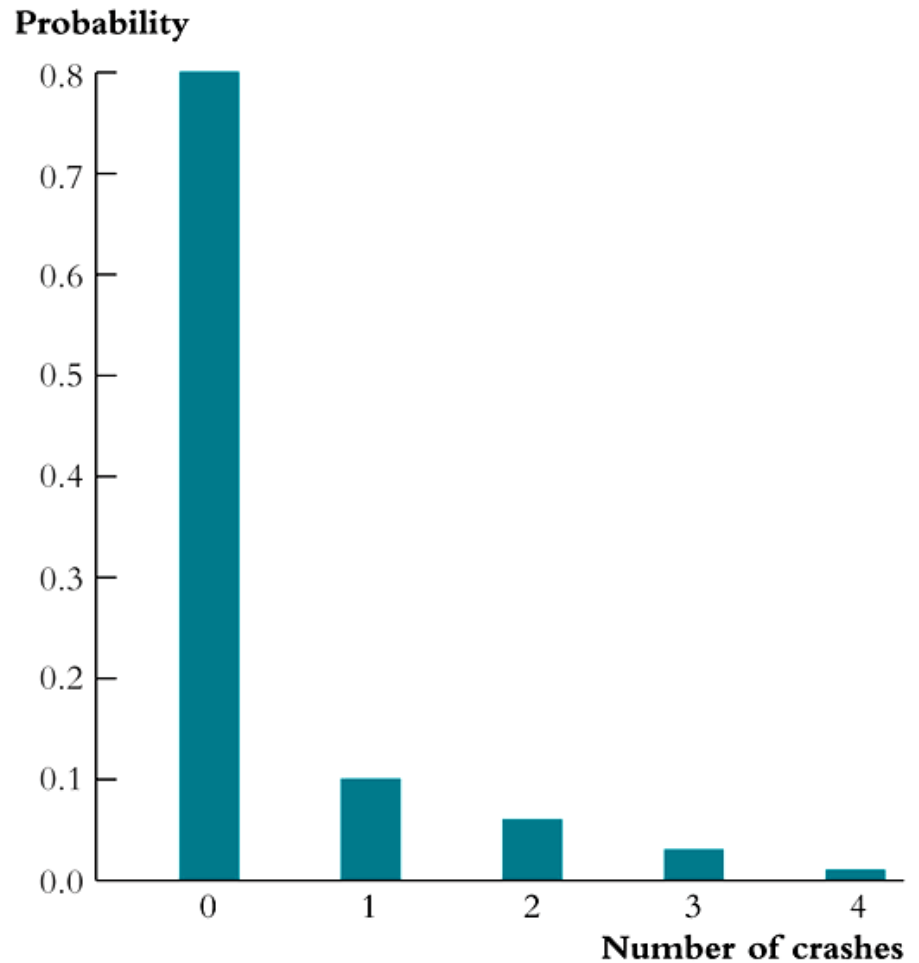
- **For a random variable X the probability distribution (or density) function of X , $f(X=a)$ is a formula giving the probability that X takes the value a .**

TABLE 2.1 Probability of Your Computer Crashing M Times

	Outcome (number of crashes)				
	0	1	2	3	4
Probability distribution	0.80	0.10	0.06	0.03	0.01
Cumulative probability distribution	0.80	0.90	0.96	0.99	1.00

FIGURE 2.1 Probability Distribution of the Number of Computer Crashes

The height of each bar is the probability that the computer crashes the indicated number of times. The height of the first bar is 0.80, so the probability of 0 computer crashes is 80%. The height of the second bar is 0.1, so the probability of 1 computer crash is 10%, and so forth for the other bars.





3. Expected Value

Expected Value and the Mean

Suppose the random variable Y takes on k possible values, y_1, \dots, y_k , where y_1 denotes the first value, y_2 denotes the second value, etc., and that the probability that Y takes on y_1 is p_1 , the probability that Y takes on y_2 is p_2 , and so forth. The expected value of Y , denoted $E(Y)$, is

$$E(Y) = y_1 p_1 + y_2 p_2 + \cdots + y_k p_k = \sum_{i=1}^k y_i p_i \quad (2.4)$$

where the notation “ $\sum_{i=1}^k y_i p_i$ ” means “the sum of $y_i p_i$ for i running from 1 to k .” The expected value of Y is also called the mean of Y or the expectation of Y and is denoted μ_Y .



Variance

Variance and Standard Deviation

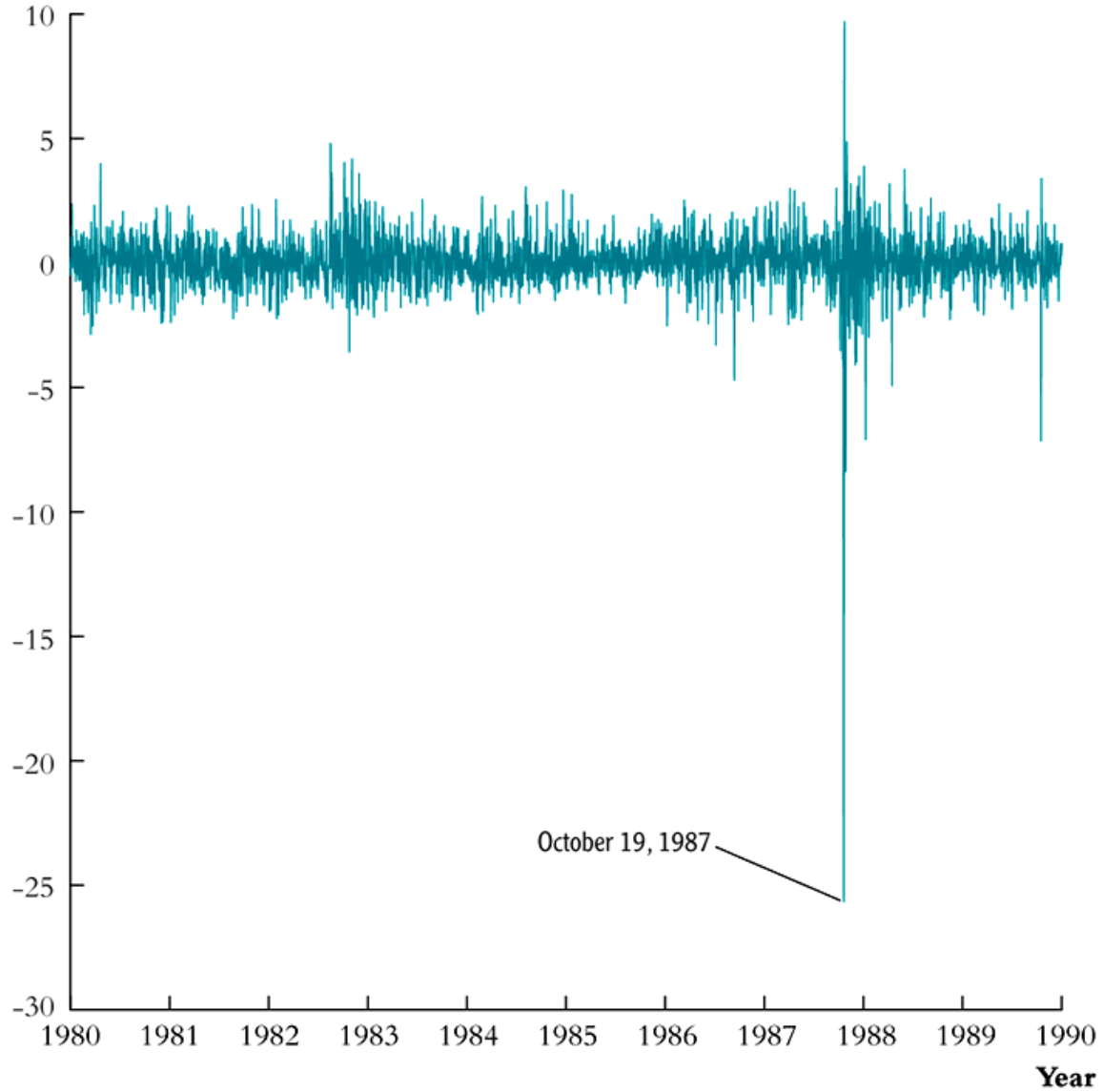
The variance of the discrete random variable Y , denoted σ_Y^2 , is

$$\sigma_Y^2 = \text{var}(Y) = E[(Y - \mu_Y)^2] = \sum_{i=1}^k (y_i - \mu_Y)^2 p_i \quad (2.6)$$

The standard deviation of Y is σ_Y , the square root of the variance. The units of the standard deviation are the same as the units of Y .

FIGURE 2.5 Daily Percentage Changes in the Dow Jones Industrial Average in the 1980s

During the 1980s, the average percentage daily change of “the Dow” index was 0.05% and its standard deviation was 1.16%. On October 19, 1987—“Black Monday”—the index fell 25.6%, or more than 22 standard deviations.





Variances of sums of RVs

Means, Variances, and Covariances of Sums of Random Variables

Let X , Y , and V be random variables, let μ_X and σ_X^2 be the mean and variance of X , let σ_{XY} be the covariance between X and Y (and so forth for the other variables), and let a , b , and c be constants. The following facts follow from the definitions of the mean, variance, and covariance:

$$E(a + bX + cY) = a + b\mu_X + c\mu_Y, \quad (2.29)$$

$$\text{var}(a + bY) = b^2\sigma_Y^2, \quad (2.30)$$

$$\text{var}(aX + bY) = a^2\sigma_X^2 + 2ab\sigma_{XY} + b^2\sigma_Y^2, \quad (2.31)$$

$$E(Y^2) = \sigma_Y^2 + \mu_Y^2, \quad (2.32)$$

$$\text{cov}(a + bX + cV, Y) = b\sigma_{XY} + c\sigma_{VY}, \text{ and} \quad (2.33)$$

$$E(XY) = \sigma_{XY} + \mu_X\mu_Y. \quad (2.34)$$

$$|\text{corr}(X, Y)| \leq 1 \text{ and } |\sigma_{XY}| \leq \sqrt{\sigma_X^2\sigma_Y^2} \text{ (correlation inequality)}. \quad (2.35)$$

$\text{Cor}(X, Y)$
 $E[(X - \mu_X)(Y - \mu_Y)]$



4. Random Sampling

Simple Random Sampling and i.i.d. Random Variables

In a simple random sample, n objects are drawn at random from a population and each object is equally likely to be drawn. The value of the random variable Y for the i^{th} randomly drawn object is denoted Y_i . Because each object is equally likely to be drawn and the distribution of Y_i is the same for all i , the random variables Y_1, \dots, Y_n are independently and identically distributed (i.i.d.); that is, the distribution of Y_i is the same for all $i = 1, \dots, n$ and Y_1 is distributed independently of Y_2, \dots, Y_n and so forth.



5. Sample size matters

Convergence in Probability, Consistency, and the Law of Large Numbers

The sample average \bar{Y} converges in probability to μ_Y (or, equivalently, \bar{Y} is consistent for μ_Y) if the probability that \bar{Y} is in the range $\mu_Y - c$ to $\mu_Y + c$ becomes arbitrarily close to one as n increases for any constant $c > 0$. This is written as $\bar{Y} \xrightarrow{p} \mu_Y$.

The law of large numbers says that if Y_i , $i = 1, \dots, n$ are independently and identically distributed with $E(Y_i) = \mu_Y$ and $\text{var}(Y_i) = \sigma_Y^2 < \infty$, then $\bar{Y} \xrightarrow{p} \mu_Y$.



7. Working with “Normal” Random Variables

Computing Probabilities Involving Normal Random Variables

Suppose Y is normally distributed with mean μ and variance σ^2 , that is, Y is distributed $N(\mu, \sigma^2)$. Then Y is standardized by subtracting its mean and dividing by its standard deviation, that is, by computing $Z = (Y - \mu) / \sigma$.

Let c_1 and c_2 denote two numbers with $c_1 < c_2$, and let $d_1 = (c_1 - \mu) / \sigma$ and $d_2 = (c_2 - \mu) / \sigma$. Then,

$$\Pr(Y \leq c_2) = \Pr(Z \leq d_2) = \Phi(d_2), \quad (2.38)$$

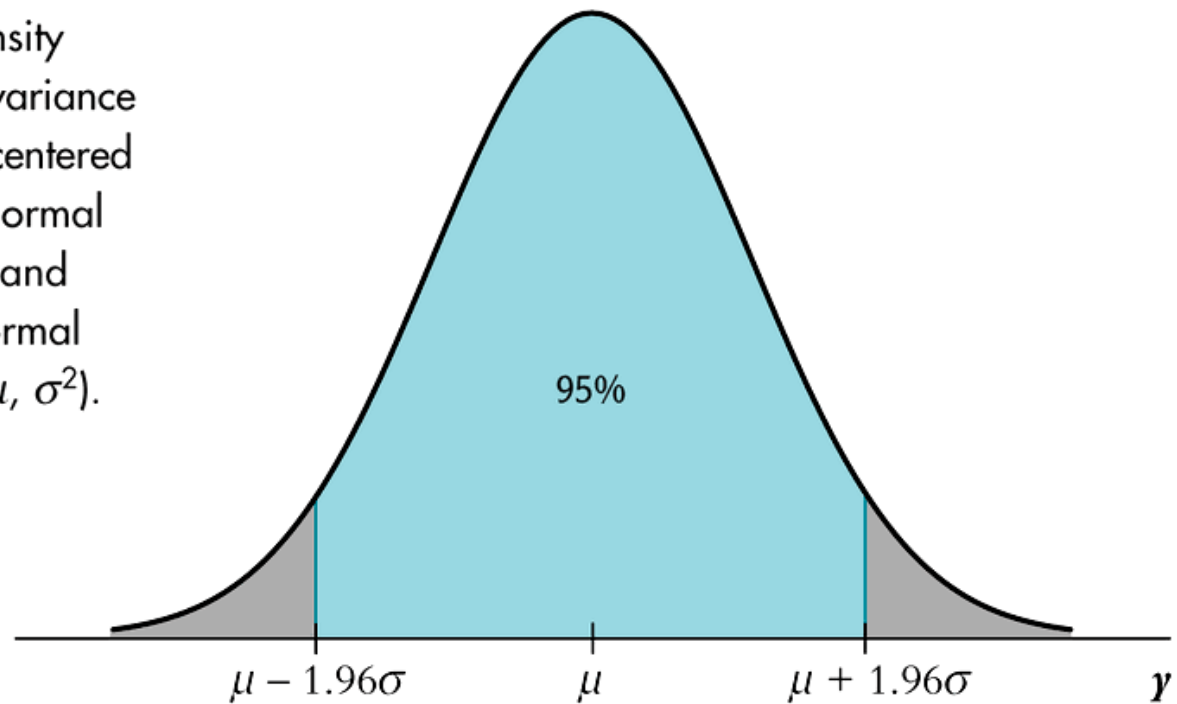
$$\Pr(Y \geq c_1) = \Pr(Z \geq d_1) = 1 - \Phi(d_1), \text{ and} \quad (2.39)$$

$$\Pr(c_1 \leq Y \leq c_2) = \Pr(d_1 \leq Z \leq d_2) = \Phi(d_2) - \Phi(d_1). \quad (2.40)$$

The normal cumulative distribution function Φ is tabulated in Appendix Table 1.

FIGURE 2.3 The Normal Probability Density

The normal probability density function with mean μ and variance σ^2 is a bell-shaped curve, centered at μ . The area under the normal p.d.f. between $\mu - 1.96\sigma$ and $\mu + 1.96\sigma$ is 0.95. The normal distribution is denoted $N(\mu, \sigma^2)$.

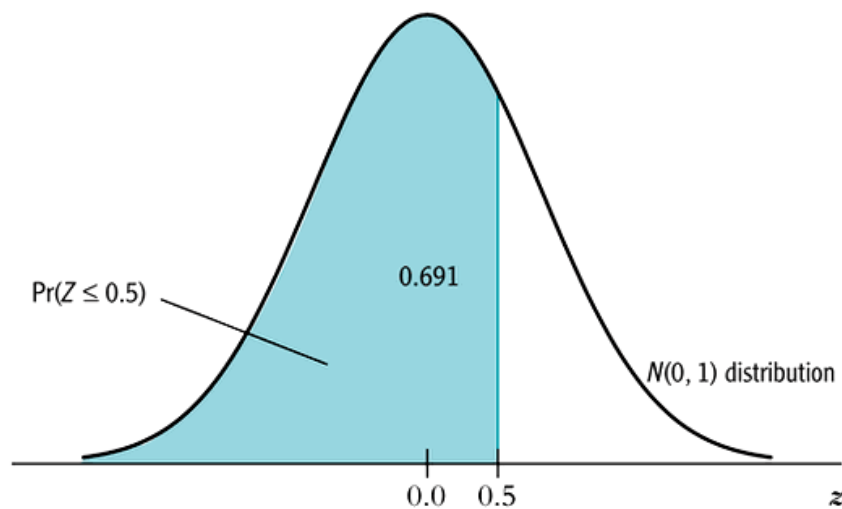
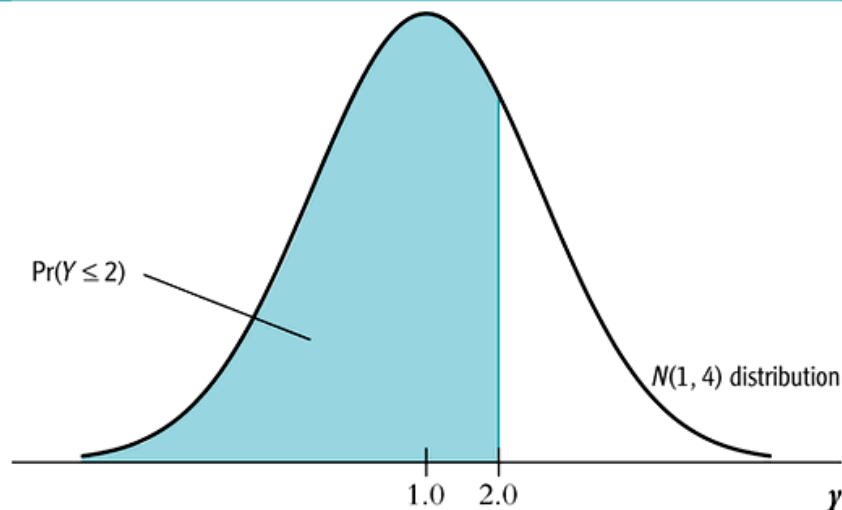


Three reasons to love the Normal Distribution are:

- 1.
- 2.
- 3.

FIGURE 2.4 Calculating the Probability that $Y \leq 2$ When Y is Distributed $N(1, 4)$

To calculate $\Pr(Y \leq 2)$, standardize Y , then use the standard normal distribution table. Y is standardized by subtracting its mean ($\mu = 1$) and dividing by its standard deviation ($\sigma_Y = 2$). The probability that $Y \leq 2$ is shown in Figure 2.4a, and the corresponding probability after standardizing Y is shown in Figure 2.4b. Because the standardized random variable, $\frac{Y-1}{2}$, is a standard normal (Z) random variable, $\Pr(Y \leq 2) = \Pr\left(\frac{Y-1}{2} \leq \frac{2-1}{2}\right) = \Pr(Z \leq 0.5)$. From Appendix Table 1, $\Pr(Z \leq 0.5) = 0.691$.



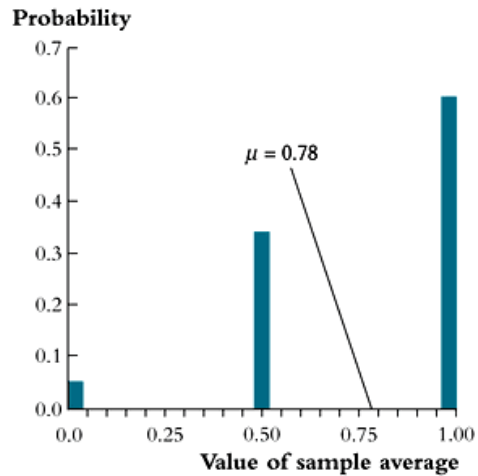


8. The Central Limit Theorem

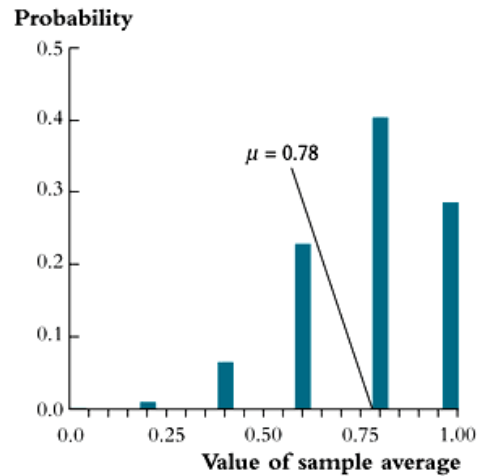
The Central Limit Theorem

Suppose that Y_1, \dots, Y_n are i.i.d. with $E(Y_i) = \mu_Y$ and $\text{var}(Y_i) = \sigma_Y^2$, where $0 < \sigma_Y^2 < \infty$. As $n \rightarrow \infty$, the distribution of $(\bar{Y} - \mu_Y) / \sigma_{\bar{Y}}$ (where $\sigma_{\bar{Y}}^2 = \sigma_Y^2 / n$) becomes arbitrarily well approximated by the standard normal distribution.

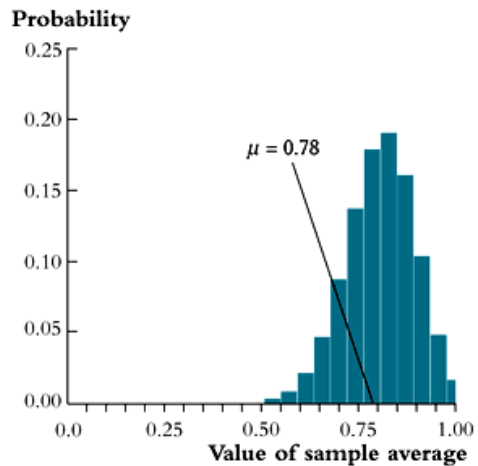
FIGURE 2.6 Sampling Distribution of the Sample Average of n Bernoulli Random Variables



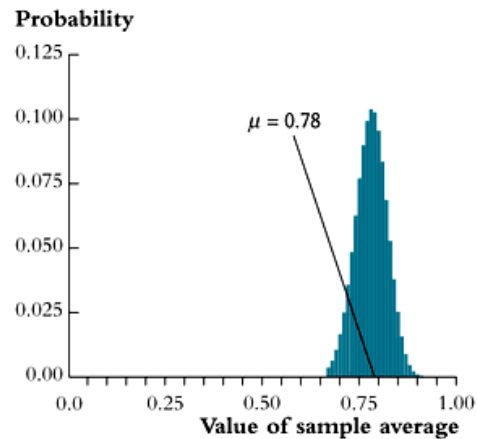
(a) $n = 2$



(b) $n = 5$



(c) $n = 25$

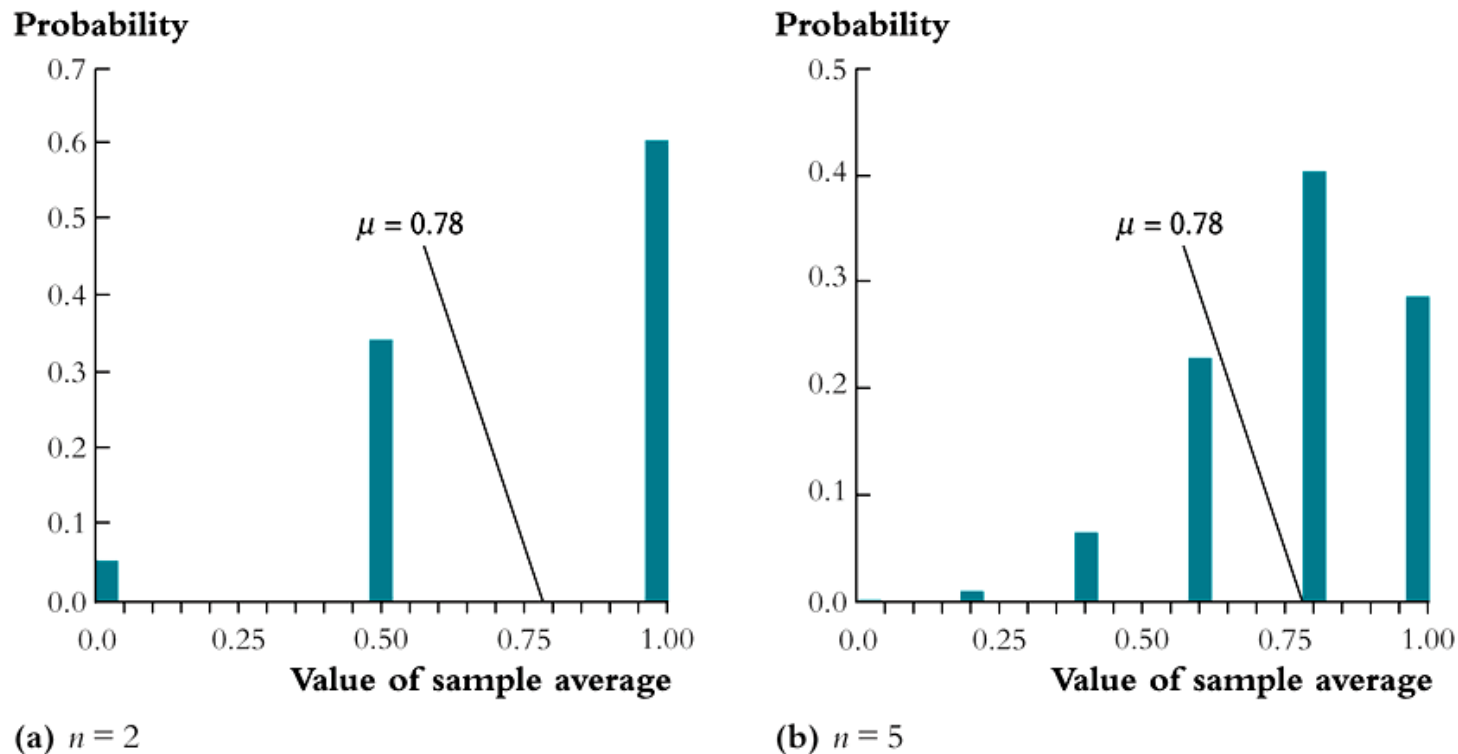


(d) $n = 100$

The distributions are the sampling distributions of \bar{Y} , the sample average of n independent Bernoulli random variables with $p = \Pr(Y_i = 1) = 0.78$ (the probability of a fast commute is 78%). The variance of the sampling distribution of \bar{Y} decreases as n gets larger, so the sampling distribution becomes more tightly concentrated around its mean $\mu = 0.78$ as the sample size n increases.

The next two slides each present one half of Figure 2.6.

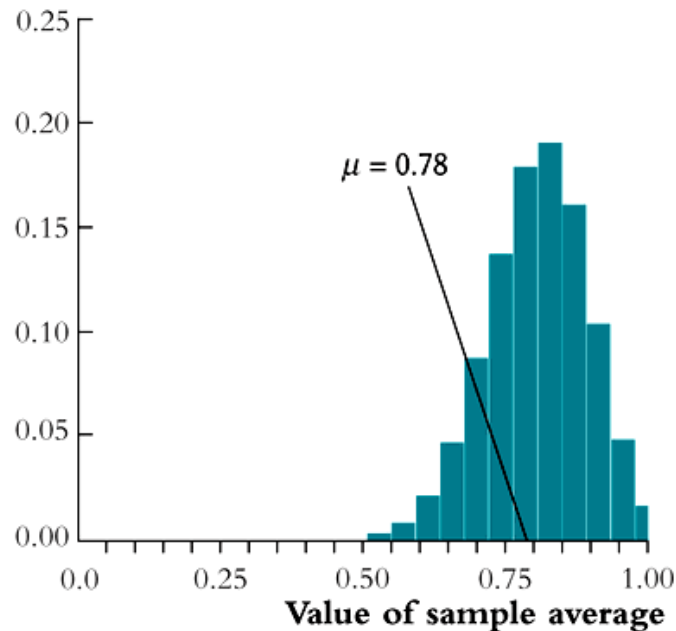
FIGURE 2.6 Sampling Distribution of the Sample Average of n Bernoulli Random Variables



The distributions are the sampling distributions of \bar{Y} , the sample average of n independent Bernoulli random variables with $p = \Pr(Y_i = 1) = 0.78$ (the probability of a fast commute is 78%). The variance of the sampling distribution of \bar{Y} decreases as n gets larger, so the sampling distribution becomes more tightly concentrated around its mean $\mu = 0.78$ as the sample size n increases.

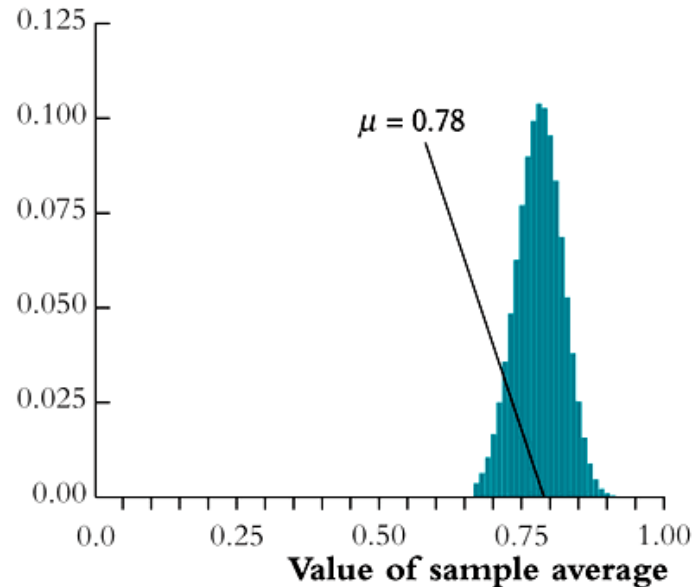
FIGURE 2.6 Sampling Distribution of the Sample Average of n Bernoulli Random Variables

Probability



(c) $n = 25$

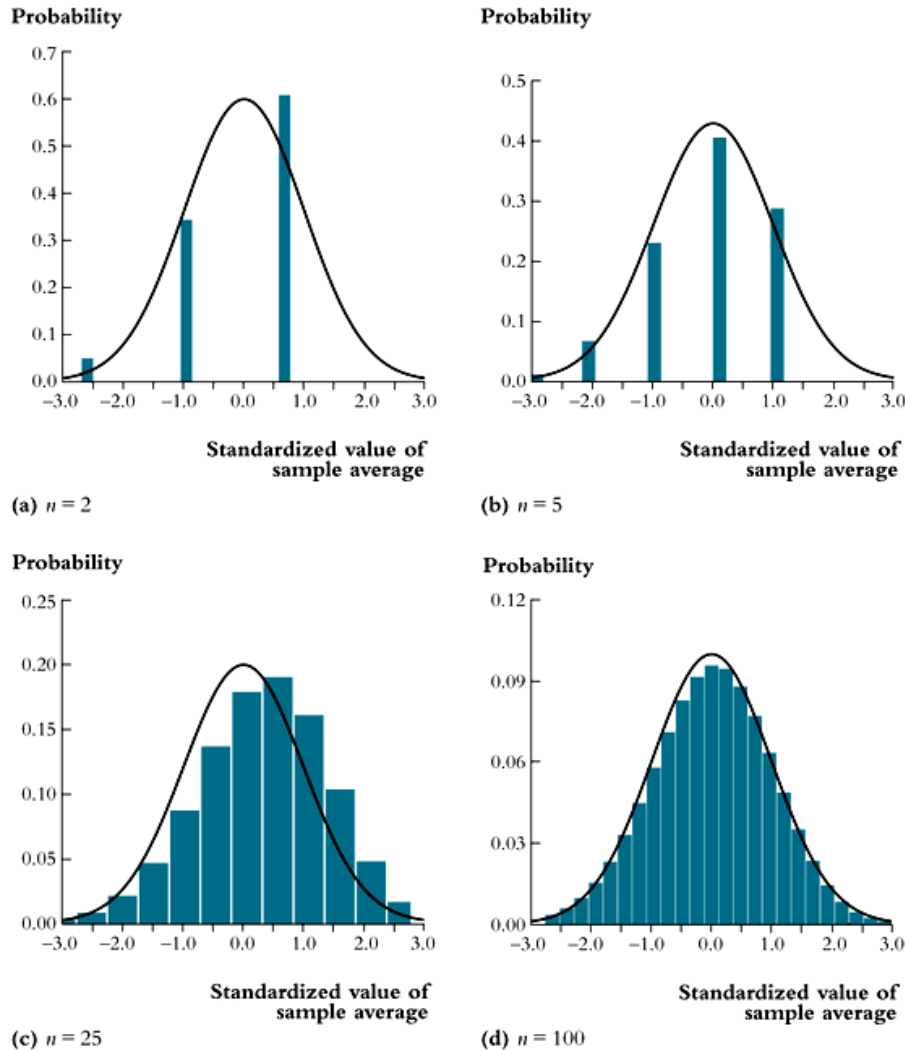
Probability



(d) $n = 100$

The distributions are the sampling distributions of \bar{Y} , the sample average of n independent Bernoulli random variables with $p = \Pr(Y_i = 1) = 0.78$ (the probability of a fast commute is 78%). The variance of the sampling distribution of \bar{Y} decreases as n gets larger, so the sampling distribution becomes more tightly concentrated around its mean $\mu = 0.78$ as the sample size n increases.

FIGURE 2.7 Distribution of the Standardized Sample Average of n Bernoulli Random Variables with $p = .78$

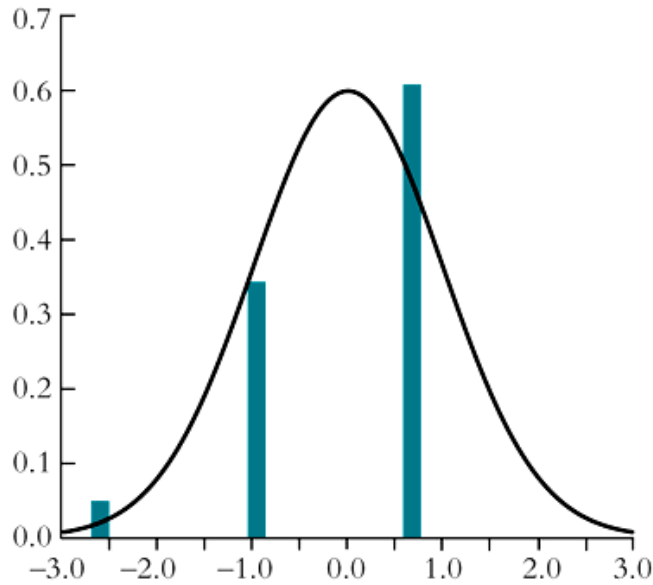


The sampling distribution of \bar{Y} in Figure 2.6 is plotted here after standardizing \bar{Y} . This centers the distributions in Figure 2.6 and magnifies the scale on the horizontal axis by a factor of \sqrt{n} . When the sample size is large, the sampling distributions are increasingly well approximated by the normal distribution (the solid line), as predicted by the central limit theorem.

The next two slides each present one half of Figure 2.7.

FIGURE 2.7 Distribution of the Standardized Sample Average of n Bernoulli Random Variables with $p = .78$

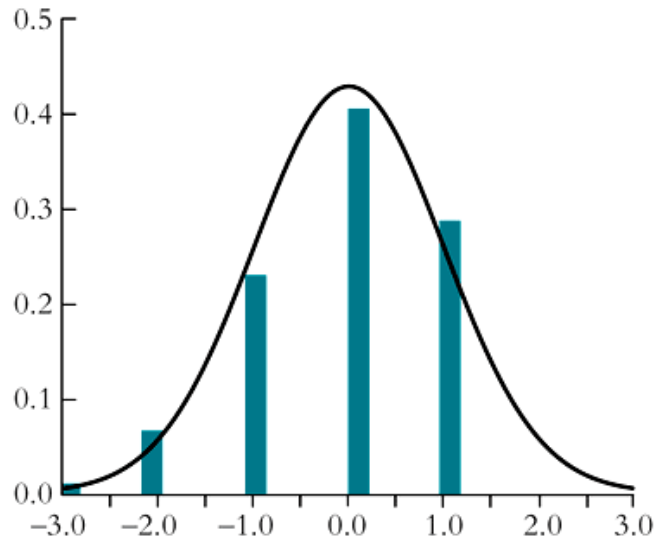
Probability



Standardized value of sample average

(a) $n = 2$

Probability



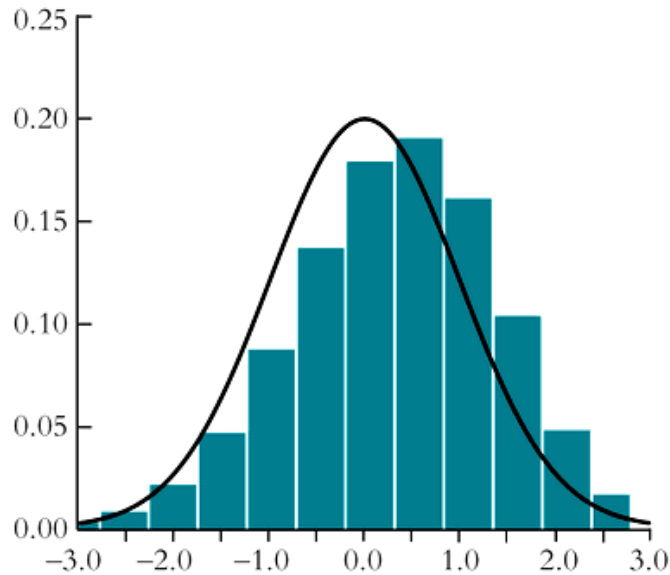
Standardized value of sample average

(b) $n = 5$

The sampling distribution of \bar{Y} in Figure 2.6 is plotted here after standardizing \bar{Y} . This centers the distributions in Figure 2.6 and magnifies the scale on the horizontal axis by a factor of \sqrt{n} . When the sample size is large, the sampling distributions are increasingly well approximated by the normal distribution (the solid line), as predicted by the central limit theorem.

FIGURE 2.7 Distribution of the Standardized Sample Average of n Bernoulli Random Variables with $p = .78$

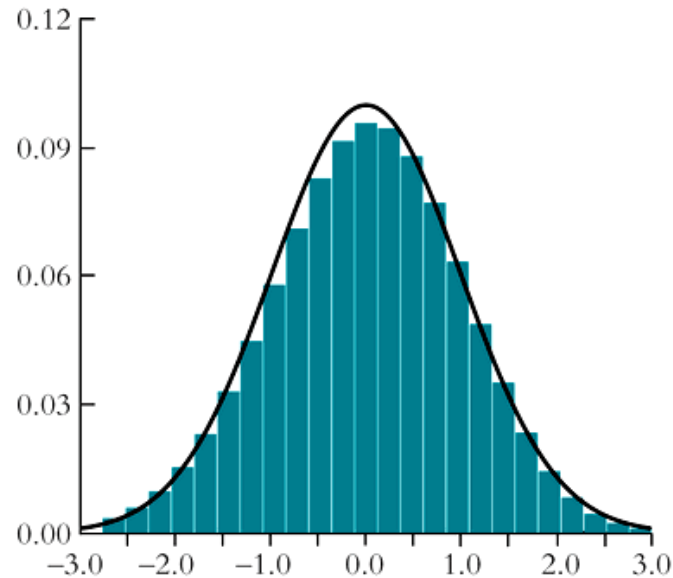
Probability



Standardized value of
sample average

(c) $n = 25$

Probability

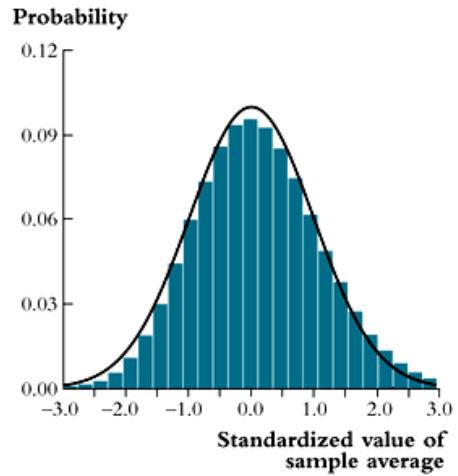
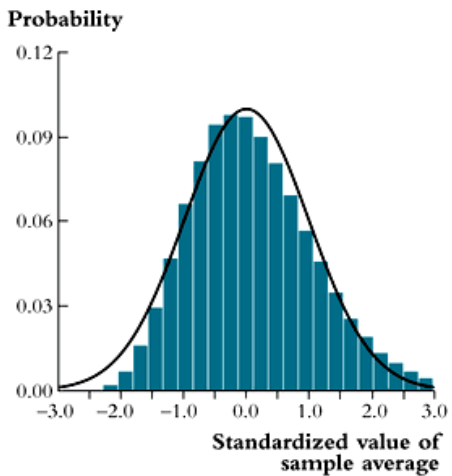
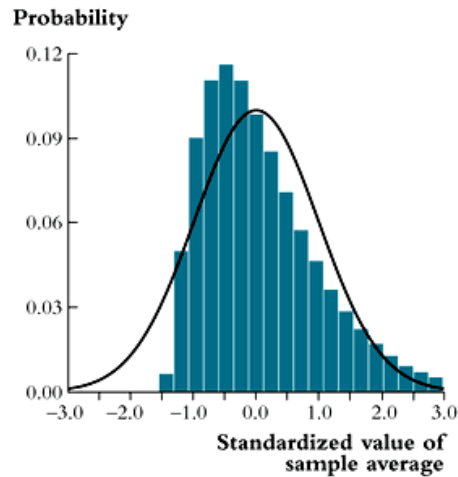
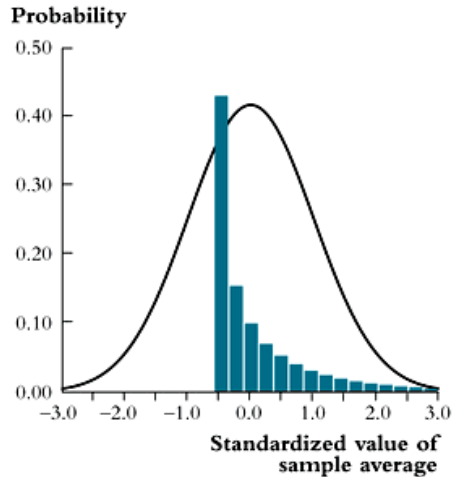


Standardized value of
sample average

(d) $n = 100$

The sampling distribution of \bar{Y} in Figure 2.6 is plotted here after standardizing \bar{Y} . This centers the distributions in Figure 2.6 and magnifies the scale on the horizontal axis by a factor of \sqrt{n} . When the sample size is large, the sampling distributions are increasingly well approximated by the normal distribution (the solid line), as predicted by the central limit theorem.

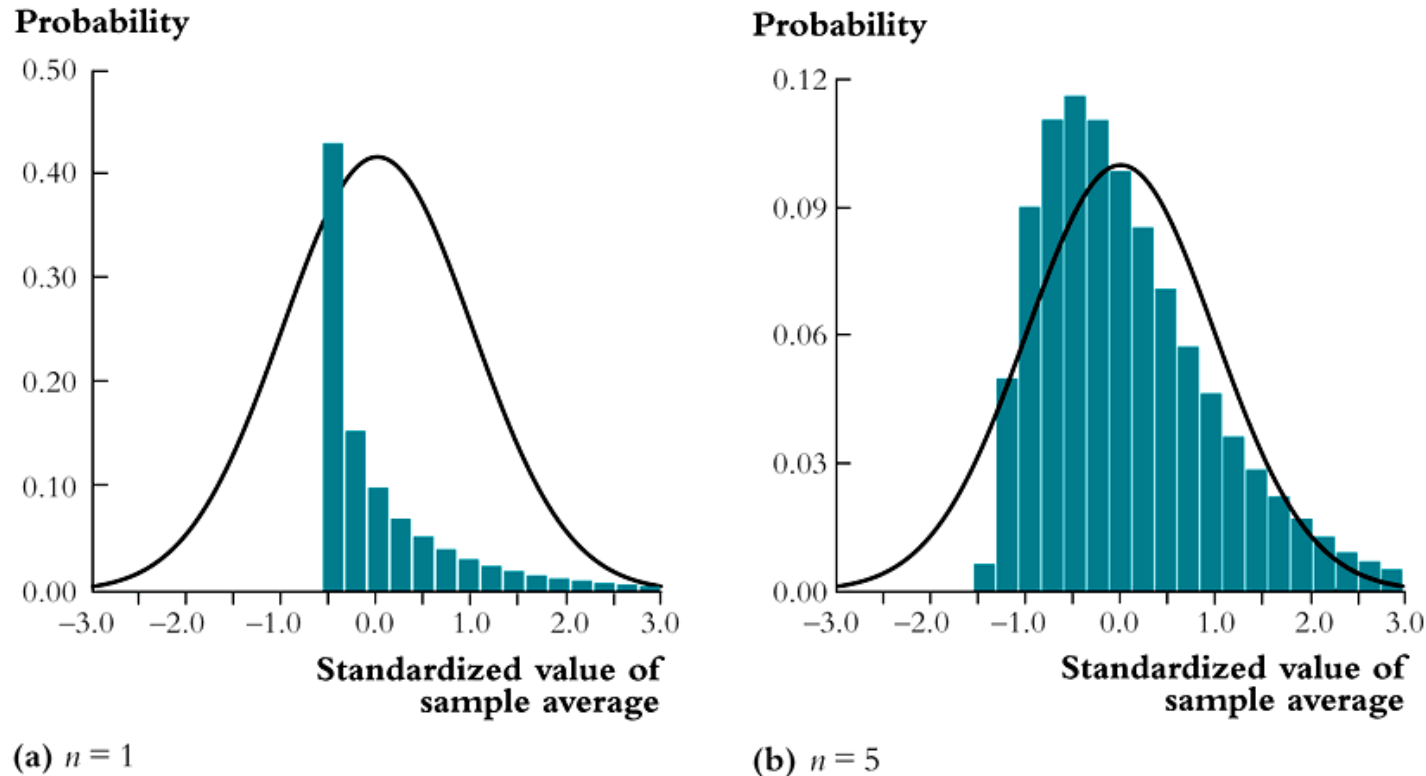
FIGURE 2.8 Distribution of the Standardized Sample Average of n Draws from a Skewed Distribution



The figures show the sampling distribution of the standardized sample average of n draws from the skewed (asymmetric) population distribution shown in Figure 2.8a. When n is small ($n = 5$), the sampling distribution, like the population distribution, is skewed. But when n is large ($n = 100$), the sampling distribution is well approximated by a standard normal distribution (solid line), as predicted by the central limit theorem.

The next two slides each present one half of Figure 2.8.

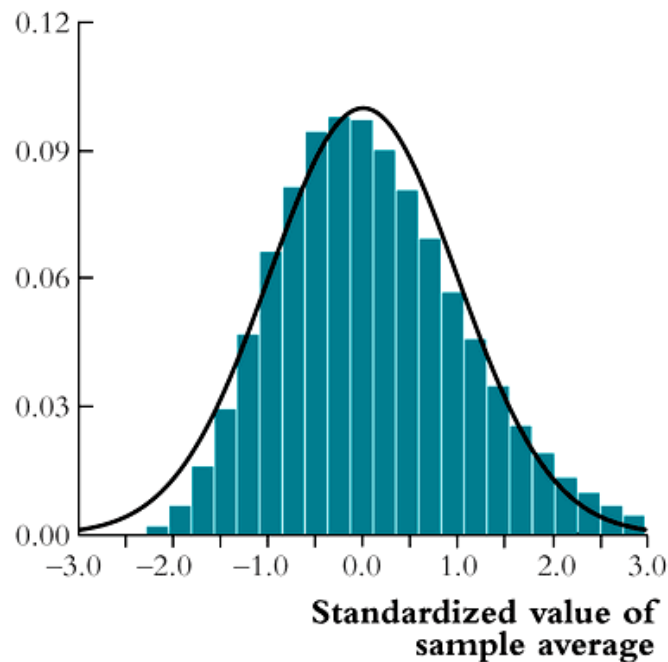
FIGURE 2.8 Distribution of the Standardized Sample Average of n Draws from a Skewed Distribution



The figures show the sampling distribution of the standardized sample average of n draws from the skewed (asymmetric) population distribution shown in Figure 2.8a. When n is small ($n = 5$), the sampling distribution, like the population distribution, is skewed. But when n is large ($n = 100$), the sampling distribution is well approximated by a standard normal distribution (solid line), as predicted by the central limit theorem.

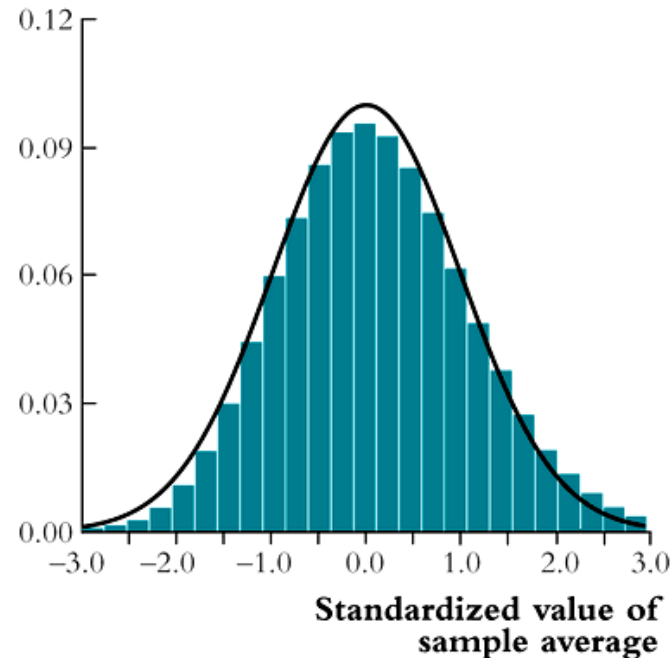
FIGURE 2.8 Distribution of the Standardized Sample Average of n Draws from a Skewed Distribution

Probability



(c) $n = 25$

Probability



(d) $n = 100$

The figures show the sampling distribution of the standardized sample average of n draws from the skewed (asymmetric) population distribution shown in Figure 2.8a. When n is small ($n = 5$), the sampling distribution, like the population distribution, is skewed. But when n is large ($n = 100$), the sampling distribution is well approximated by a standard normal distribution (solid line), as predicted by the central limit theorem.

Econometrics 120B

Summary of Probability Lecture

1. **Probability**
2. **Random Variables**
3. **Expected Value**
4. **Variance, Covariance**
5. **Random Sampling**
6. **Sample size matters**
7. **Working with Normal Distribution**
8. **Central Limit Theorem**



Next Class: Statistics Review

Problem Set #1

Problem Set #1
Due Thursday, January 24

Please answer the questions on this sheet (or photocopy) in pen.

1. Polling

Hillary and Barack compete in the primary of a large state. They are the only two candidates. Five minutes after the polls close the Constant News Network announces the results of an exit poll. In a random, representative sample of 200 primary voters (all of whom tell the truth), they find that 56% voted for Hilary.

Let p be the true proportion of primary voters that voted for Hillary.

- a. Calculate and report a 95% confidence interval for p , showing your steps.

Problem Set #1: Data Gathering

2. Data gathering exercise. (*Submit a page or two with two graphs, stapled to this question sheet.*)

a. Gather data from a *controlled experiment*. Report the values of one variable (X) that you controlled (or randomly assigned) and another (Y) that was influenced by X. You should have at least 15 observations. (The “demand for coffee” survey performed in class was an experiment in which X was price and Y was quantity demanded.) *Do not* survey people on their demand for coffee.

Report the source of the data, the sample mean of each variable and attach an X-Y graph (scatterplot). (Stata is good at this. Recall that Stata is available in the lab or can be leased on the Internet.)