Probability owes its existence to gambling

- It is remarkable that a science, which began with the consideration of games of chance, should be elevated to the rank of the most important subject of human knowledge. —Pierre Simon Laplace

- “一門開始於研究賭博的科學, 竟然成為人類知識中最重要的學科, 這無疑是令人驚訝的事情。”

Probability and AI

- Why does probability help in developing automatic reasoning?

  Answer 1. Logic programming is defective

  Answer 2. From the bionic viewpoint, it is not a bad idea to learn from human brains, is it?

  Pros: Our brain works like a statistician

  Cons: Ok, let’s have an experience!

Assessing Probability

- There are three viewpoints about how to assess the probability of an uncertain event:

    1. Classical probability

    2. Relative frequency probability

    3. Subjective probability

Classical Probability

- Gerolamo Cardano 1550 (1663)
  Count the number of equally possible outcomes, the proportion relating to an event.
- The spirit of a priori probability:
  1. Count the number of all possible outcomes
  2. Attach equally likely probability to each
  3. Count the number of outcomes for an event

$$P = \frac{\#Event}{\#Sample\ Space}$$

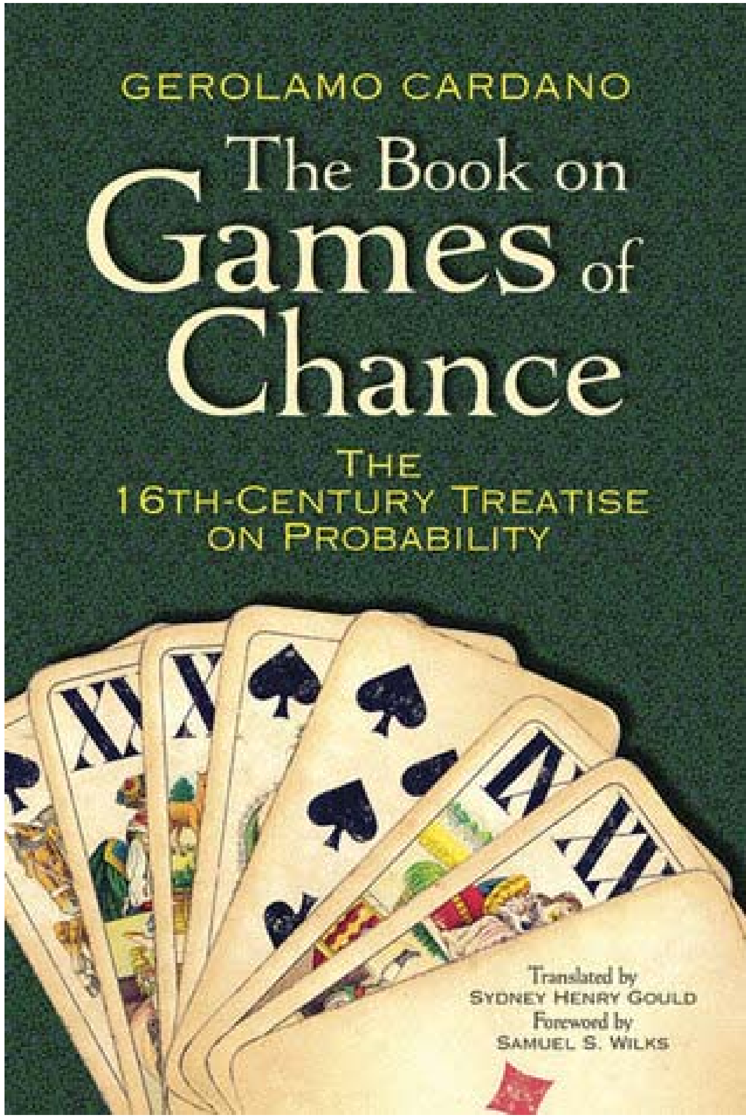
De Vetula by de Fournival (13’ century)

3	18	Punctatura	1	Cadentia	1
4	17	Punctatura	1	Cadentiæ	3
5	16	Punctaturæ	2	Cadentiæ	6
6	15	Punctaturæ	3	Cadentiæ	10
7	14	Punctaturæ	4	Cadentiæ	15
8	13	Punctaturæ	5	Cadentiæ	21
9	12	Punctaturæ	6	Cadentiæ	25
10	11	Punctaturæ	6	Cadentiæ	27

16th Centry

- Galilieo: Fair dice and sample space

- Gerolamo Cardano 1550 (1663)
  Exclusive events: sum
  Independent events: multiply



17th Centry

- Pascal and Fermat's correspondence (1654 July)
  De Mere: "At least on 6 in 4 rolls" vs "at least on double-six in 24 throws of 2 dice" (crucial)

  Truth is the same of Toulouse and at Paris. --by Pascal

- Huygens' first printed text on probability (1656)
  Defines the expectation

Assessing Probability

- There are three viewpoints about how to assess the probability of an uncertain event:

1. Classical probability

2. **Relative frequency probability**

3. Subjective probability

Assessing Probability

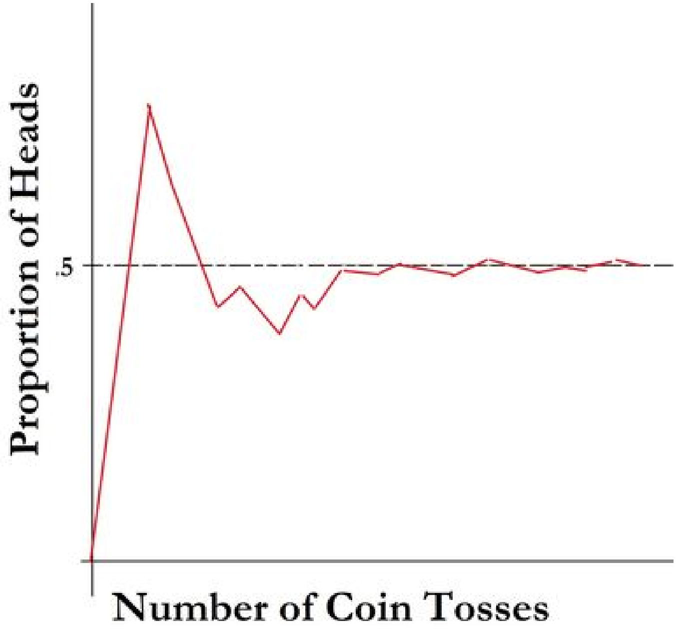
2. Relative-Frequency Probability (客觀的機率理論)
長期重複試行的隨機實驗中，事件的相對次數會漸趨穩定。而此一次數的比率即為該事件的機率。

$$P(E) = \lim_{n \to \infty} \frac{n(E)}{n}$$

即資料越多越接近真實值

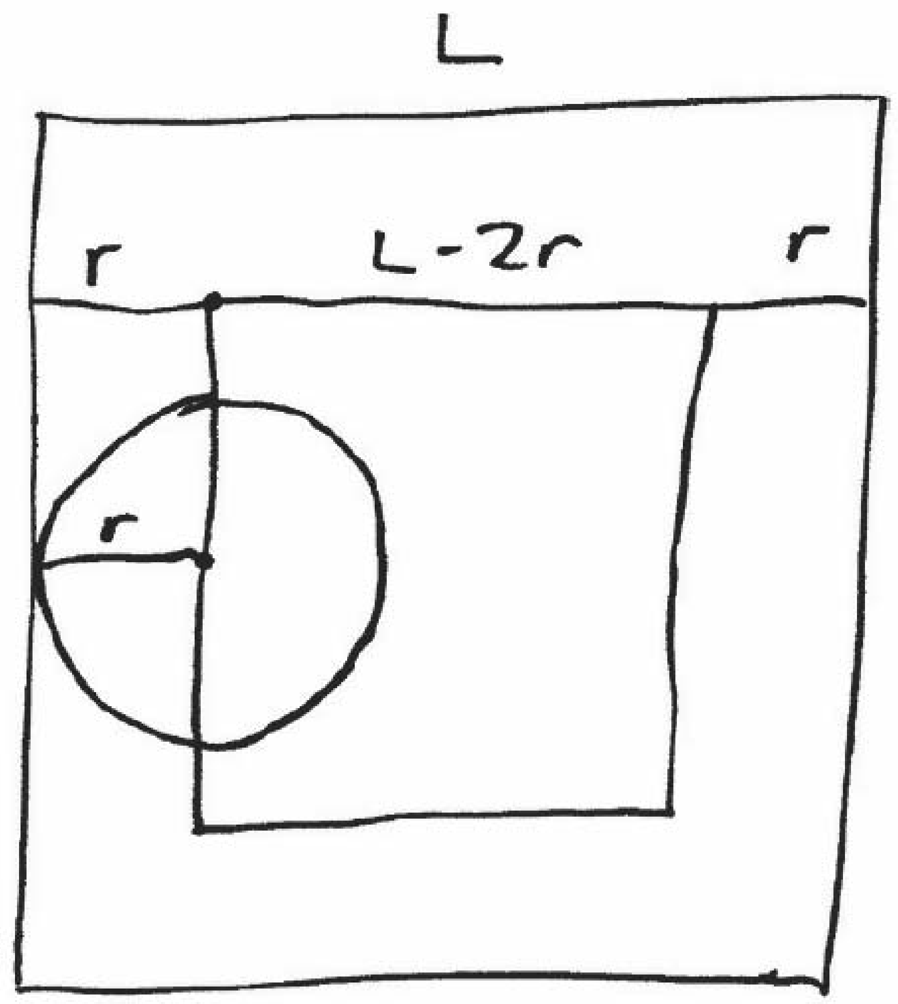
Relative Frequency Probability

- 若某事件有既定的機率，而我們不斷的進行相同的實驗，則該事件發生的次數比率會越來越接近這個既定的機率。

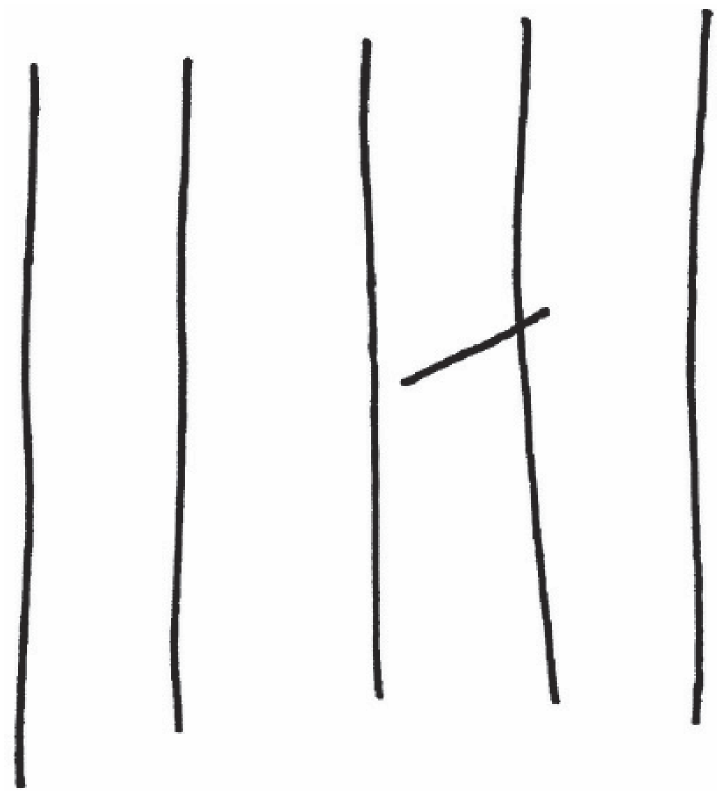


Franc-Carreau

Probability and Geometry

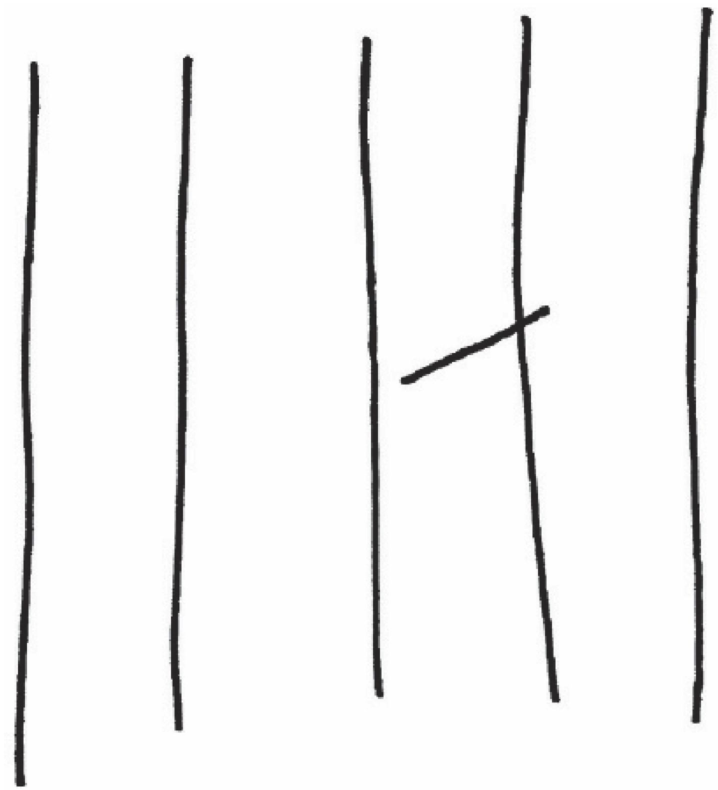


Buffon’s Needle Problem



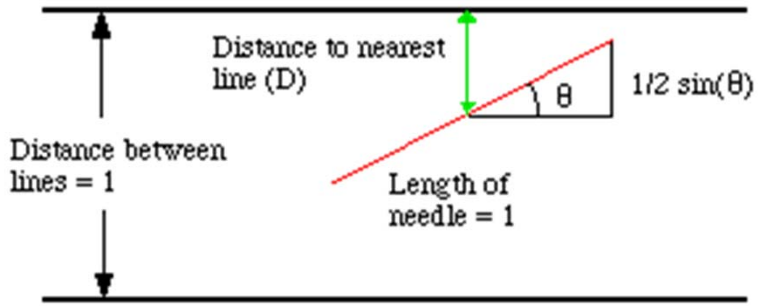
Buffon’s Needle Problem

- Georges-Louis Leclerc, Comte de Buffon (1777)



Buffon's Needle Problem

- Buffon's method: Proceed the experiment with a large number of times (1777)
- The needle will hit the line if the closest distance to a line (D) is less than or equal to 1/2 times the sine of theta. That is, $D \leq (1/2)\sin(\theta)$.

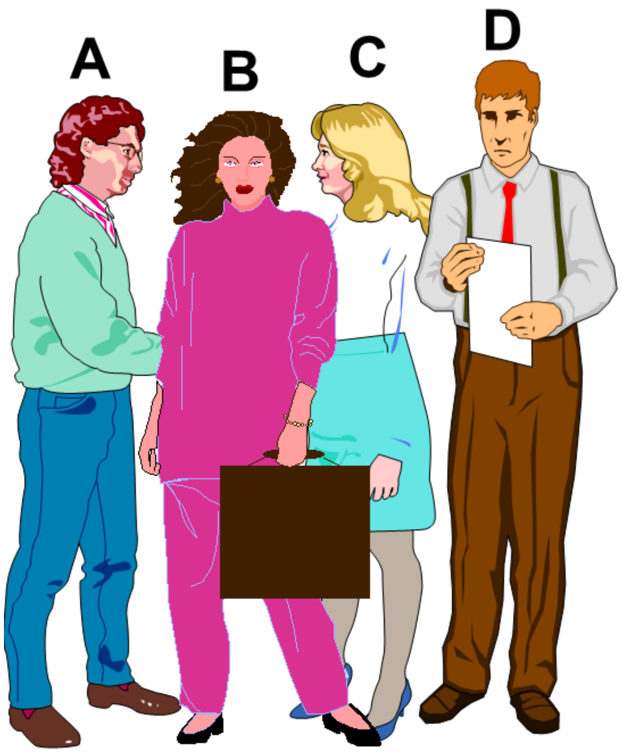


Sampling Distributions

- A sampling distribution is a probability distribution of all of the possible values of a statistic for a given size sample selected from a population

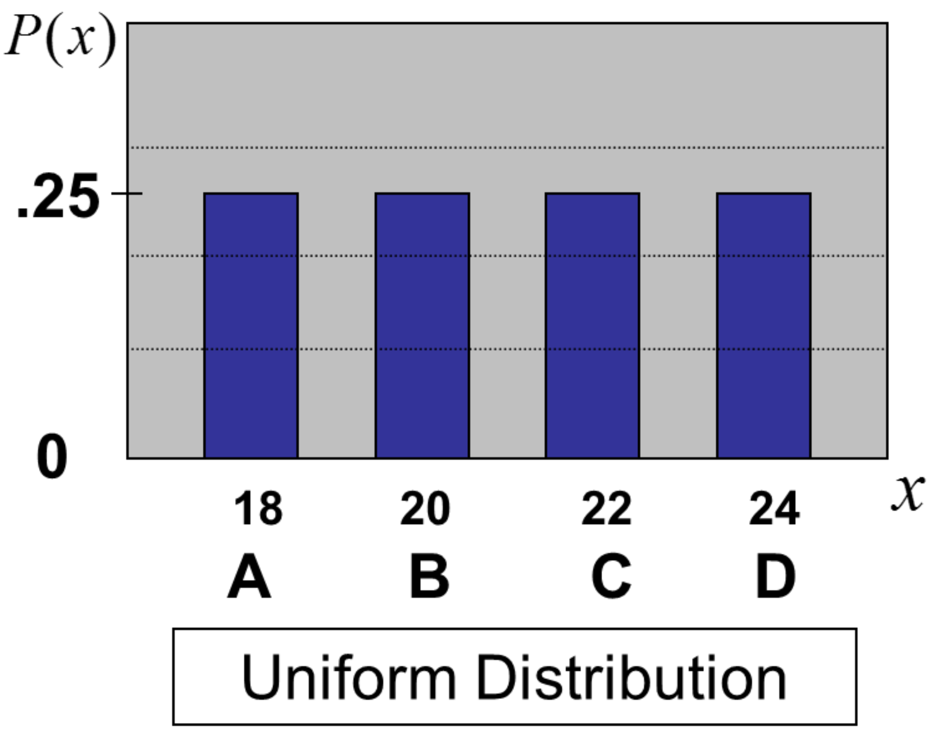
Developing a Sampling Distribution (1 of 6)

- **Assume there is a population …**
- Population size $N=4$
- Random variable, X, is age of individuals
- Values of X: 18, 20, 22, 24 (years)



Developing a Sampling Distribution (2 of 6)

In this example the Population Distribution is uniform:



Developing a Sampling Distribution (3 of 6)

Now consider all possible samples of size *n* = 2

1st Obs	2nd Observation 18	20	22	24
18	18,18	18,20	18,22	18,24
20	20,18	20,20	20,22	20,24
22	22,18	22,20	22,22	22,24
24	24,18	24,20	24,22	24,24

16 possible samples (sampling with replacement)

16 Sample Means

1st Obs	2nd Observation 18	20	22	24
18	18	19	20	21
20	19	20	21	22
22	20	21	22	23
24	21	22	23	24

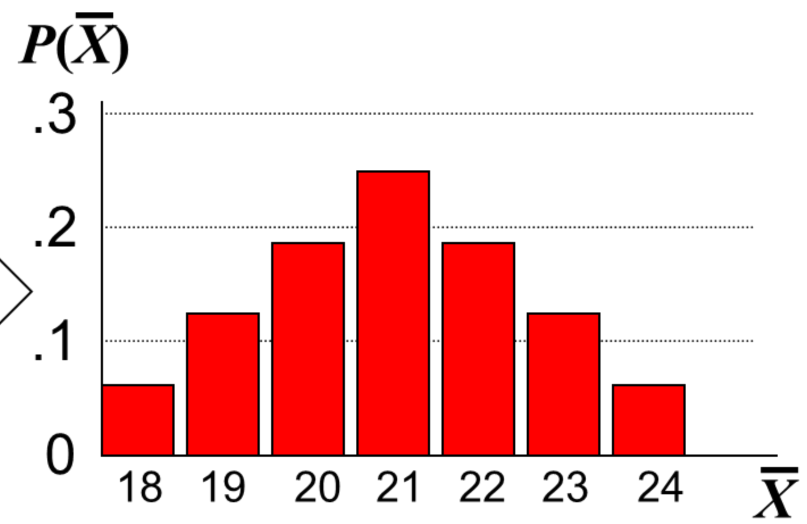
Developing a Sampling Distribution (4 of 6)

Sampling Distribution of All Sample Means

16 Sample Means

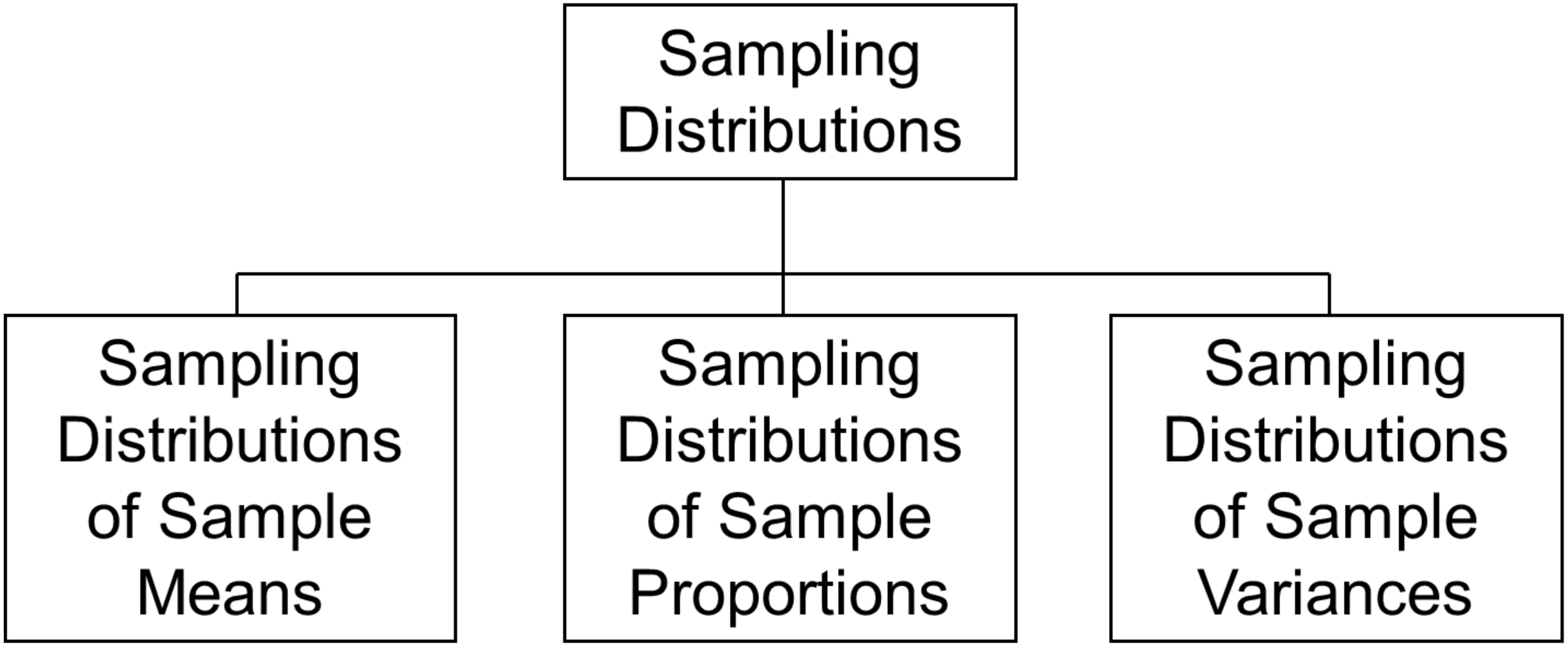
1st Obs	2nd Observation			
	18	20	22	24
18	18	19	20	21
20	19	20	21	22
22	20	21	22	23
24	21	22	23	24

Distribution of Sample Means

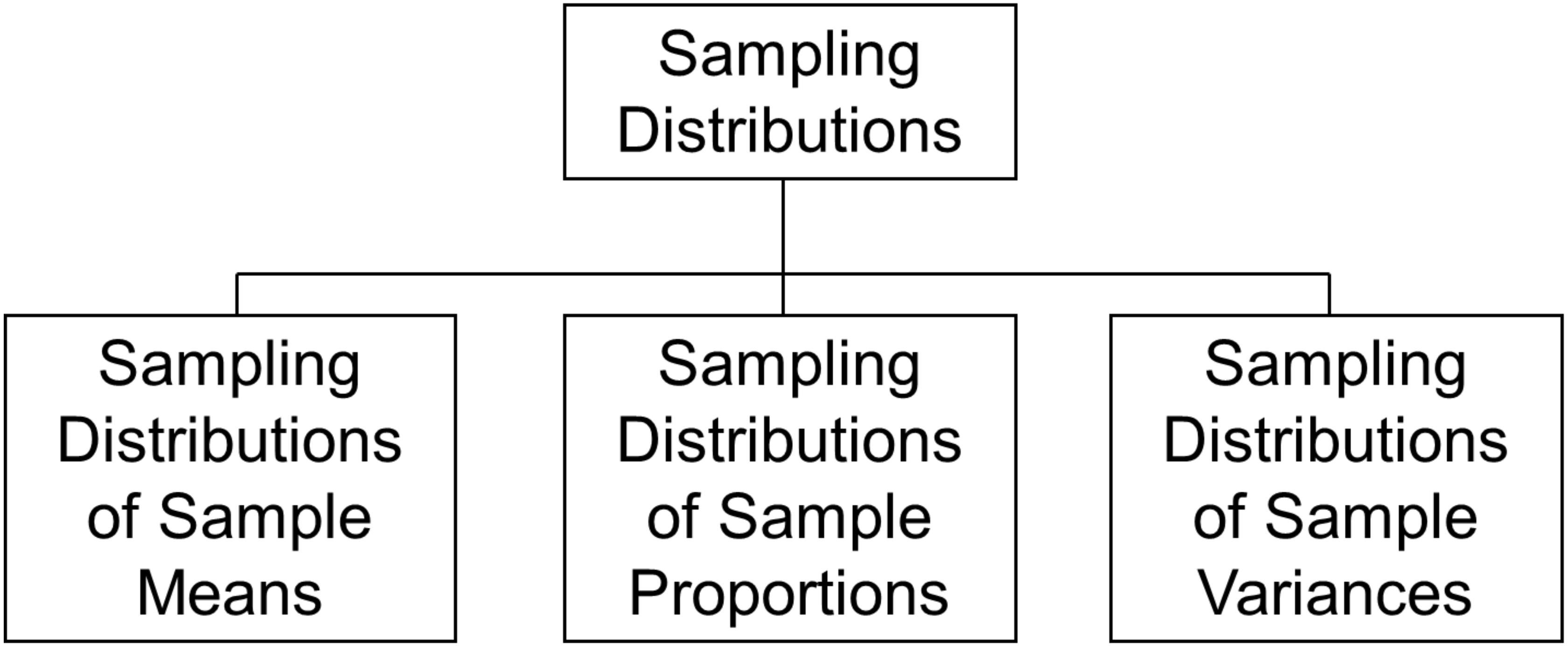


(no longer uniform)

Chapter Outline



Section 6.2 Sampling Distributions of Sample Means



Sample Mean

- Let $X_1, X_2, \ldots, X_n$ represent a random sample from a population
- The sample mean value of these observations is defined as

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Standard Error of the Mean

- Different samples of the same size from the same population will yield different sample means
- A measure of the variability in the mean from sample to sample is given by the Standard Error of the Mean:

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

- Note that the standard error of the mean decreases as the sample size increases

Comparing the Population with Its Sampling Distribution

Population

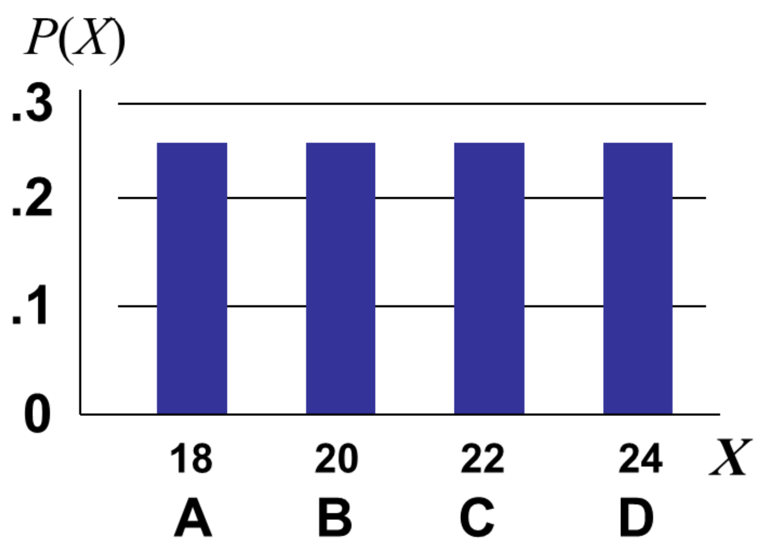
$N = 4$

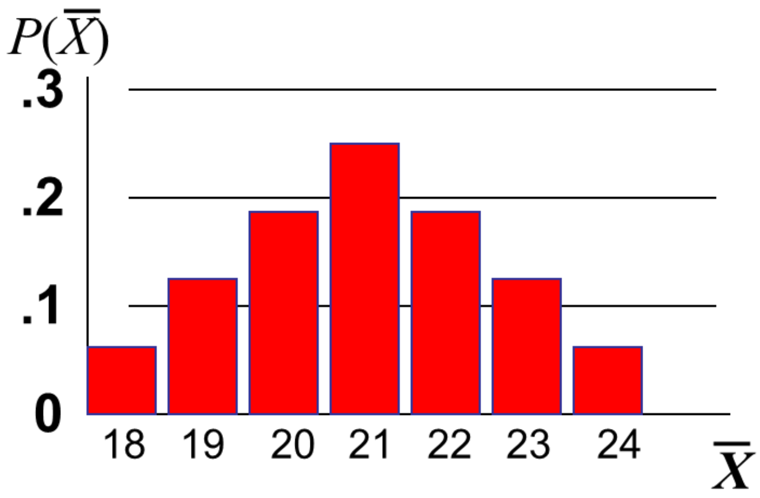
$\mu = 21 \quad \sigma = 2.236$

Sample Means Distribution

$n = 2$

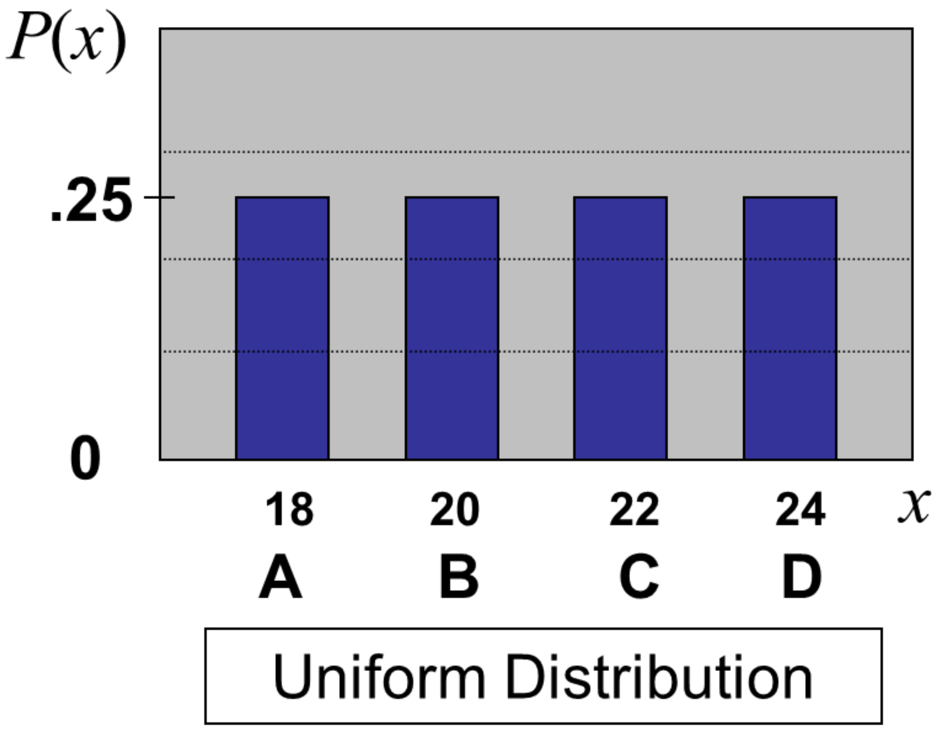
$\mu_{\bar{X}} = 21 \quad \sigma_{\bar{X}} = 1.58$





Developing a Sampling Distribution (5 of 6)

Summary Measures for the Population Distribution:



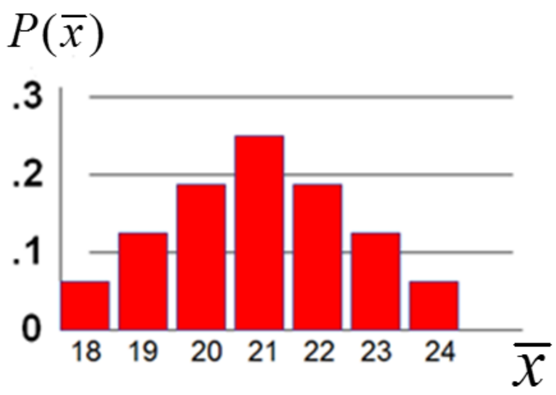
$$\mu = \frac{\sum X_i}{N}$$

$$= \frac{18 + 20 + 22 + 24}{4} = 21$$

$$\sigma = \sqrt{\frac{\sum (X_i - \mu)^2}{N}} = 2.236$$

Developing a Sampling Distribution (6 of 6)

Summary Measures of the Sampling Distribution:



$$E(\bar{X}) = \frac{\sum \bar{X}_i}{N} = \frac{18 + 19 + 21 + \cdots + 24}{16} = 21 = \mu$$

$$\sigma_{\bar{x}} = \sqrt{\frac{\sum (\bar{X}_i - \mu)^2}{N}}$$

$$= \sqrt{\frac{(18-21)^2 + (19-21)^2 + \cdots + (24-21)^2}{16}} = 1.58$$

Standard Normal Distribution for the Sample Means

- Z-value for the sampling distribution of $\bar{X}$:

$$z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

where: $\bar{X}$ = sample mean

μ = population mean

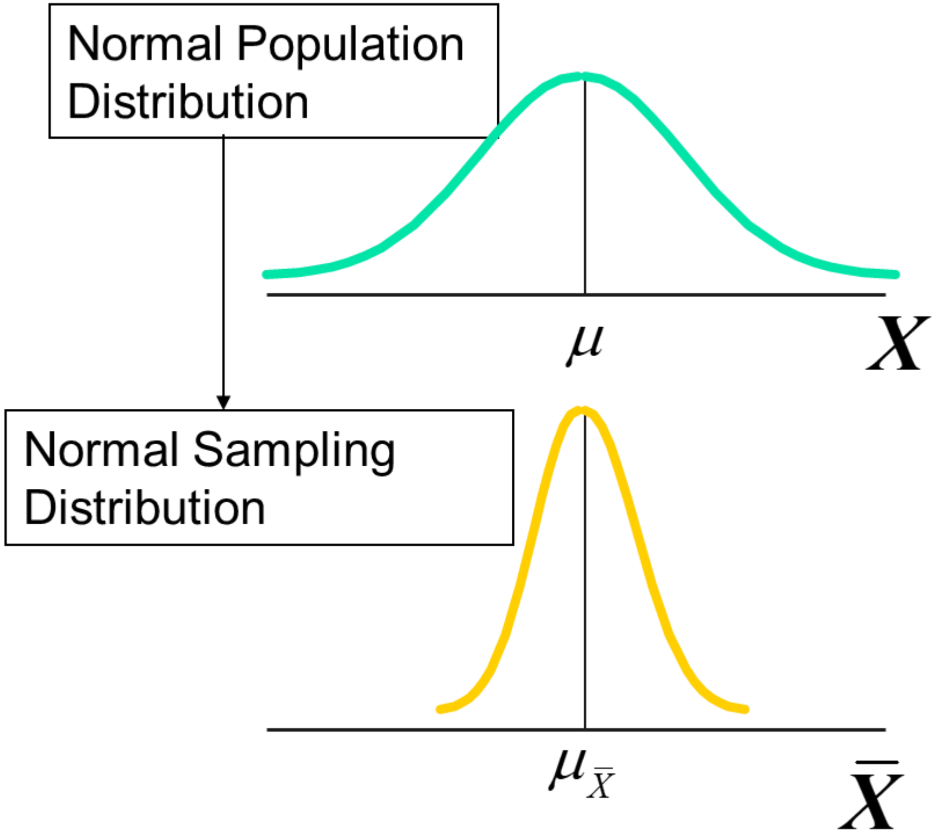
$\sigma_{\bar{X}}$ = standard error of the mean

Z is a standardized normal random variable with mean of 0 and a variance of 1

Sampling Distribution Properties (1 of 3)

$$E\left[\bar{X}\right] = \mu$$

(i.e. $\bar{X}$ is unbiased)

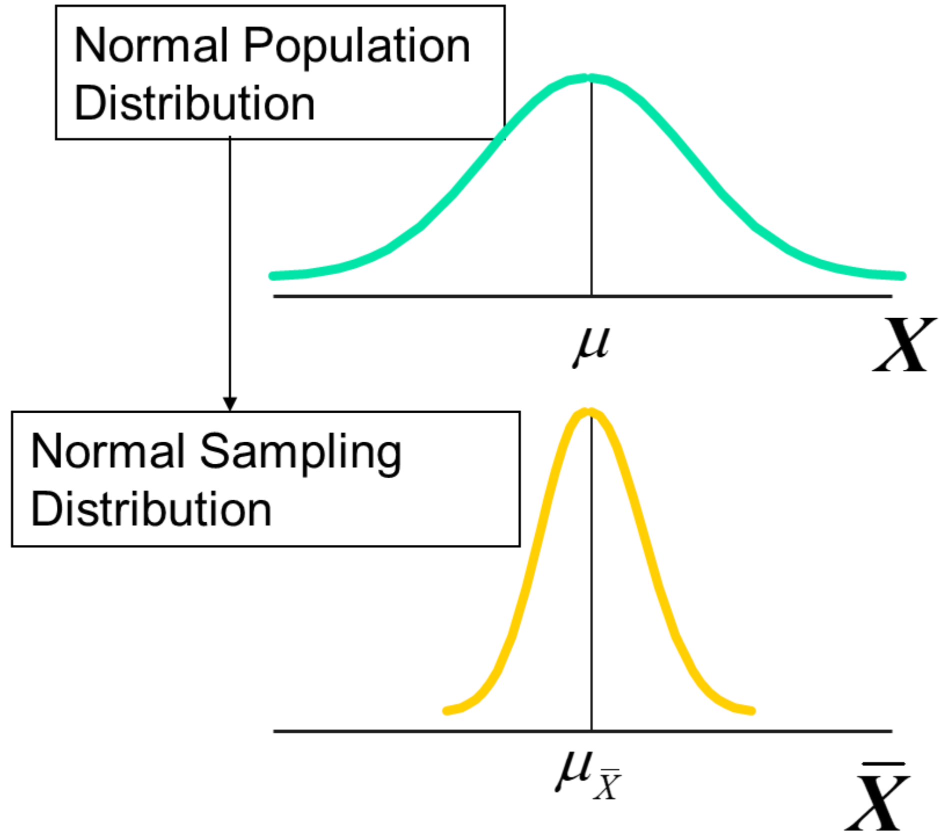


(both distributions have the same mean)

Sampling Distribution Properties (2 of 3)

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

(i.e. $\bar{X}$ is unbiased)

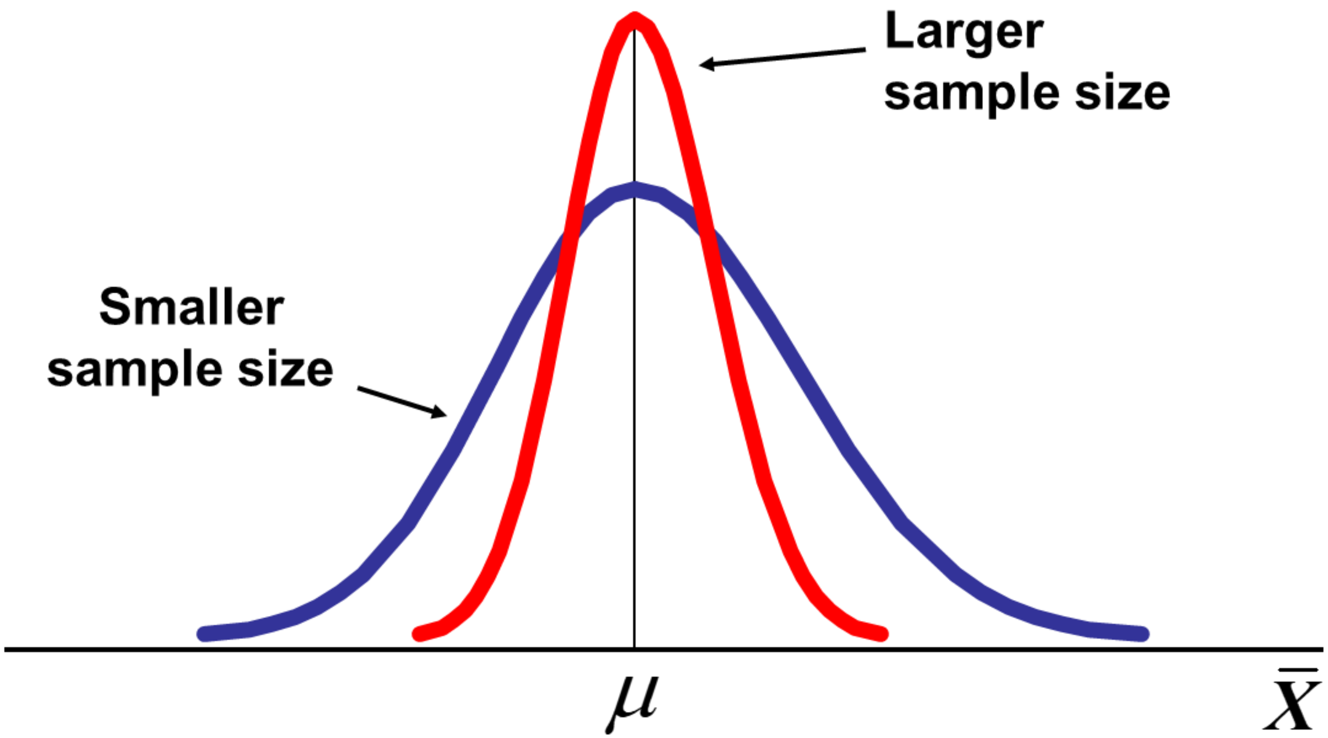


(the distribution of $\bar{X}$
has a reduced standard deviation)

Sampling Distribution Properties (3 of 3)

As *n* increases,

$\sigma_{\bar{X}}$ decreases



Central Limit Theorem (1 of 3)

- Even if the population is not normal,
- …sample means from the population will be approximately normal as long as the sample size is large enough.

Central Limit Theorem (2 of 3)

- Let $X_1, X_2, \ldots, X_n$ be a set of *n* independent random variables having identical distributions with mean μ, variance σ^2, and $\bar{X}$ as the mean of these random variables.
- As *n* becomes large, the central limit theorem states that the distribution of

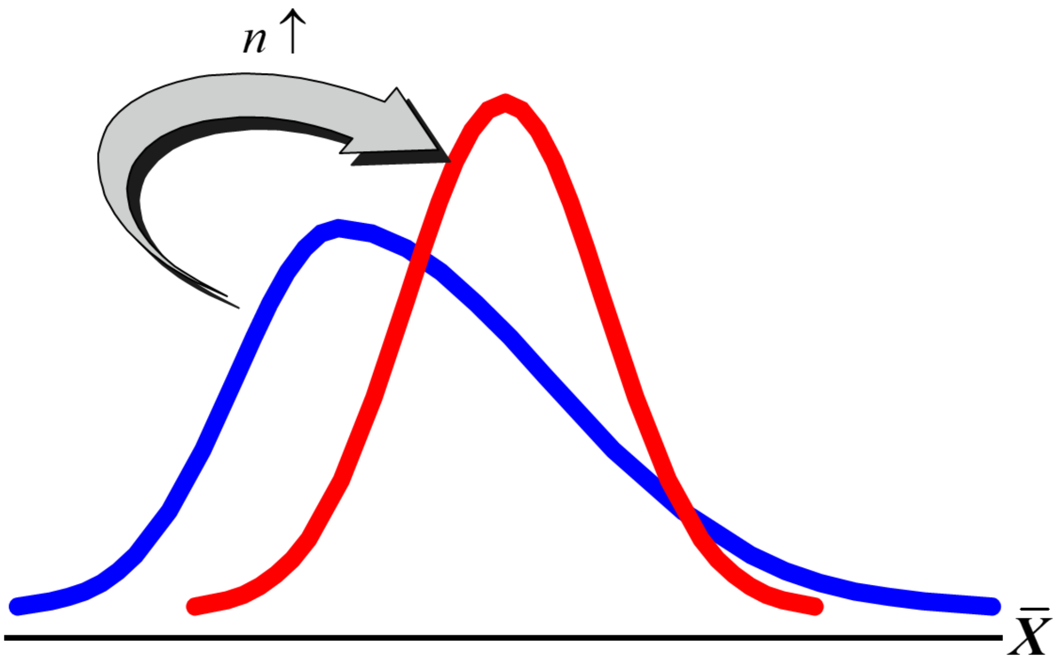
$$Z = \frac{\bar{X} - \mu_x}{\sigma_{\bar{X}}}$$

approaches the standard normal distribution

Central Limit Theorem (3 of 3)

As the sample size gets large enough…

the sampling distribution becomes almost normal regardless of shape of population



If the Population Is Not Normal

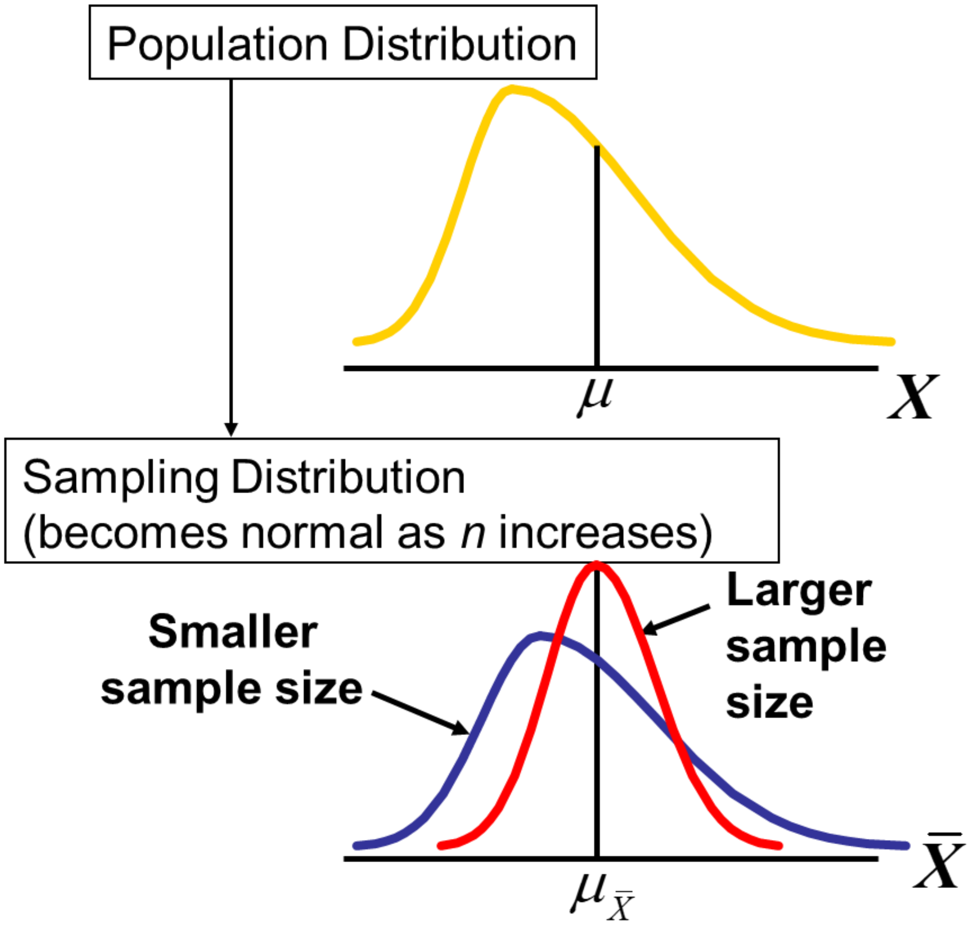
Sampling distribution properties:

Central Tendency

$$\mu_{\bar{X}} = \mu$$

Variation

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$



How Large Is Large Enough?

- For most distributions, $n > 25$ will give a sampling distribution that is nearly normal
- For normal population distributions, the sampling distribution of the mean is always normally distributed

Example 1 (1 of 3)

- Suppose a large population has mean $\mu = 8$ and standard deviation $\sigma = 3$. suppose a random sample of size n = 36 is selected.
- What is the probability that the sample mean is between 7.8 and 8.2?

Example 1 (2 of 3)

Solution:

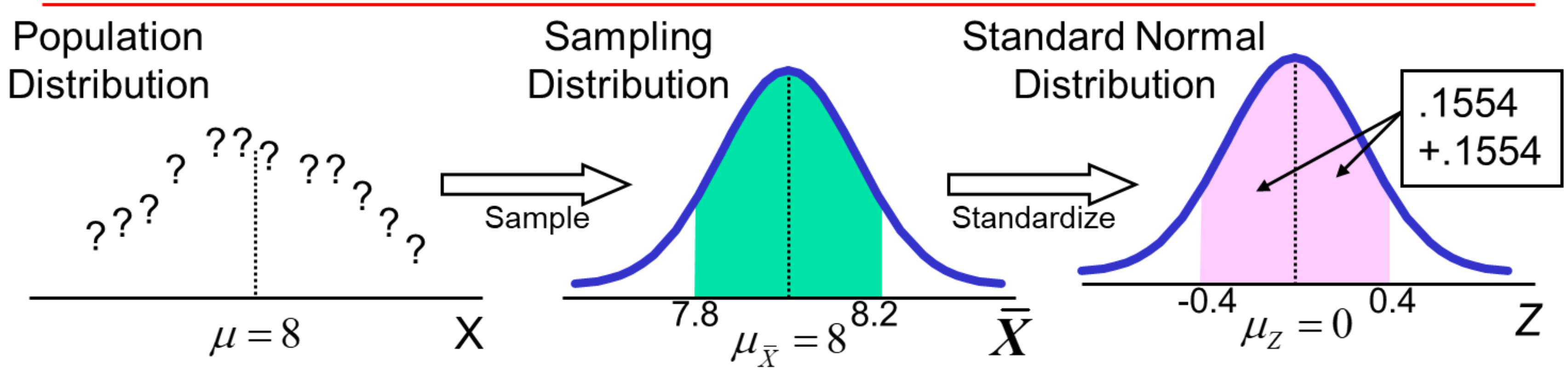
- Even if the population is not normally distributed, the central limit theorem can be used ($n > 25$)
- … so the sampling distribution of $\bar{X}$ is approximately normal
- … with mean $\mu_{\bar{X}} = 8$
- …and standard deviation $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{3}{\sqrt{36}} = 0.5$

Example 1 (3 of 3)

Solution: (continued):

$$P(7.8 < \mu_{\bar{x}} < 8.2) = P\left(\frac{7.8-8}{\frac{3}{\sqrt{36}}} < \frac{\mu_{\bar{x}} - \mu}{\frac{\sigma}{\sqrt{n}}} < \frac{8.2-8}{\frac{3}{\sqrt{36}}} \right)$$

$$= P(-0.4 < Z < 0.4) = \boxed{0.3108}$$



Acceptance Intervals

- Goal: determine a range within which sample means are likely to occur, given a population mean and variance
  - By the Central Limit Theorem, we know that the distribution of $\bar{X}$ is approximately normal if *n* is large enough, with mean μ and standard deviation $\sigma_{\bar{X}}$
  - Let $z_{\frac{\alpha}{2}}$ be the *z*-value that leaves area $\frac{\alpha}{2}$ in the upper tail of the normal distribution (i.e., the interval $-z_{\frac{\alpha}{2}}$ to $z_{\frac{\alpha}{2}}$ encloses probability $1-\alpha$)
  - Then

$$\mu \pm z_{\frac{\alpha}{2}} \sigma_{\bar{X}}$$

is the interval that includes $\bar{X}$ with probability $1-\alpha$