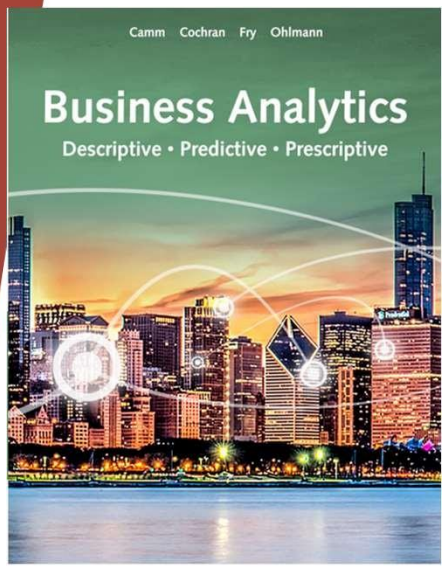


Business Analytics

Chapter 6 Statistical Inference

HAMLINE UNIVERSITY SCHOOL OF BUSINESS

4

Statistical Inference

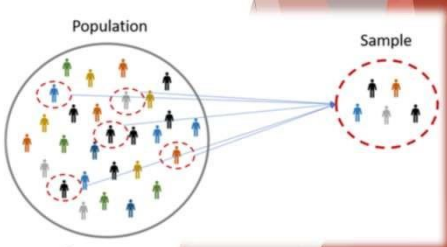
If we want to know something about a population, why can't we just ask everyone?

- We can:
  - With - A Census
    - collects data from every element in the population of interest
- We can't always ask everyone though:
  - Expensive.
  - Time consuming.
  - Misleading.
  - Unnecessary.
  - Impractical.

5

Statistical Inference

- **Statistical inference**
  - Uses sample data to make estimates of or draw conclusions about one or more characteristics of a population.
- The **sampled population**
  - is the population from which the sample is drawn.
- A **frame**
  - is a list of elements from which the sample will be selected.



6

Sampling Examples

- Election Year!
  - Texas members of a political party are considering supporting a candidate
  - Need to estimate the proportion of voters that favor the candidate in Texas
  - Of 400 voters, 160 prefer the candidate = 40%

- New Tires
  - Tire company makes a new tire that is supposed to have increased life on the roads
  - Makes 120 tires for testing
  - Mean of 36,500 miles

7

Selecting a Sample

Sampling from a Finite Population

Sampling from an Infinite Population

8

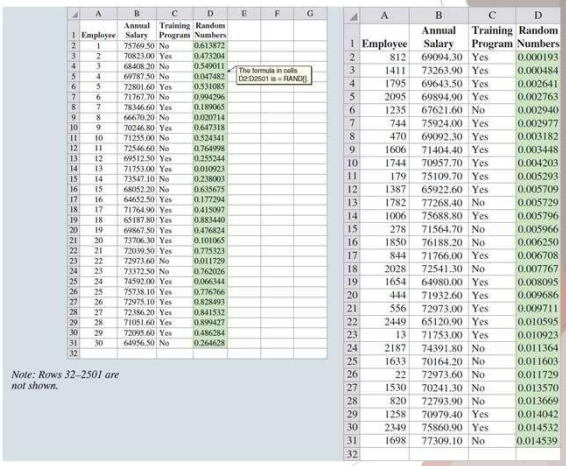
Selecting a Sample

- **Parameter**:
  - A measurable factor that defines a characteristic of a population, process, or system.
    - Examples: Population: Mean, Standard Deviation, Correlation, Variance
- Sampling from a Finite Population:
  - make valid statistical inferences about the population.
- Simple Random Sample (Finite Population):
  - A **simple random sample** of size *n* from a finite population of size *N*
  - is a sample selected such that each possible sample of size *n* has the same probability of being selected.

9

Selecting a Sample

Figure 6.1: Using Excel to Select a Simple Random Sample



	A	B	C	D	E	F	G
1	Employee	Annual Salary	Training Program	Random Numbers			
2	1	75769.50	No	0.613872			
3	2	70823.00	Yes	0.473204			
4	3	68408.20	No	0.549011			
5	4	69787.50	No	0.047482			
6	5	72801.60	Yes	0.531085			
7	6	71767.70	No	0.994296			
8	7	78346.60	Yes	0.189065			
9	8	66670.20	No	0.020714			
10	9	70246.80	Yes	0.647318			
11	10	71255.00	No	0.524341			
12	11	72546.60	No	0.764998			
13	12	69512.50	Yes	0.255244			
14	13	71753.00	Yes	0.010923			
15	14	73547.10	No	0.238003			
16	15	68052.20	No	0.635675			
17	16	64652.50	Yes	0.177294			
18	17	71764.90	Yes	0.415097			
19	18	65187.80	Yes	0.883440			
20	19	69867.50	Yes	0.476824			
21	20	73706.30	Yes	0.101065			
22	21	72039.50	Yes	0.775323			
23	22	72973.60	No	0.011729			
24	23	73372.50	No	0.762026			
25	24	74592.00	Yes	0.066344			
26	25	75738.10	Yes	0.776766			
27	26	72975.10	Yes	0.828493			
28	27	72386.20	Yes	0.841532			
29	28	71051.60	Yes	0.899427			
30	29	72095.60	Yes	0.486284			
31	30	64956.50	No	0.264628			
32							

The formula in cells D2:D2501 is = RAND()

Note: Rows 32–2501 are not shown.

	A	B	C	D
1	Employee	Annual Salary	Training Program	Random Numbers
2	812	69094.30	Yes	0.000193
3	1411	73263.90	Yes	0.000484
4	1795	69643.50	Yes	0.002641
5	2095	69894.90	Yes	0.002763
6	1235	67621.60	No	0.002940
7	744	75924.00	Yes	0.002977
8	470	69092.30	Yes	0.003182
9	1606	71404.40	Yes	0.003448
10	1744	70957.70	Yes	0.004203
11	179	75109.70	Yes	0.005293
12	1387	65922.60	Yes	0.005709
13	1782	77268.40	No	0.005729
14	1006	75688.80	Yes	0.005796
15	278	71564.70	No	0.005966
16	1850	76188.20	No	0.006250
17	844	71766.00	Yes	0.006708
18	2028	72541.30	No	0.007767
19	1654	64980.00	Yes	0.008095
20	444	71932.60	Yes	0.009686
21	556	72973.00	Yes	0.009711
22	2449	65120.90	Yes	0.010595
23	13	71753.00	Yes	0.010923
24	2187	74391.80	No	0.011364
25	1633	70164.20	No	0.011603
26	22	72973.60	No	0.011729
27	1530	70241.30	No	0.013570
28	820	72793.90	No	0.013669
29	1258	70979.40	Yes	0.014042
30	2349	75860.90	Yes	0.014532
31	1698	77309.10	No	0.014539
32				

10

Selecting a Sample

- Sampling from an Infinite Population:
  - With an infinite population, you cannot select a simple random sample because you cannot construct a frame consisting of all the elements.
  - Statisticians recommend selecting what is called a random sample.
  - Usually associated with a process that operates over time.
    - An ongoing process
- Random Sample (Infinite Population):
  - A **random sample** of size *n* from an infinite population is a sample selected such that the following conditions are satisfied:
    1. Each element selected comes from the same population.
    2. Each element is selected independently.

11

Point Estimation

12

Point Estimation

- Sample **statistic**
  - An estimate for a population parameter

	Population	Sample
Mean	μ	x
Standard Deviation	σ	s
Proportion	p	p

The sample mean x is the **point estimator** of the population mean μ

The sample standard deviation s is **the point estimator** of the population standard deviation σ.

The sample proportion p is the **point estimator** of the population proportion p.

The numerical value obtained for x, s, or p is called the **point estimate.**

13

Point Estimation

Annual Salary and Training Program Status for a Simple Random Sample of 30 EAI Employees (Excel)

Annual Salary ($)	Management Training Program	Annual Salary ($)	Management Training Program
$x_1 = 69,094.30$	Yes	$x_{16} = 71,766.00$	Yes
$x_2 = 73,263.90$	Yes	$x_{17} = 72,541.30$	No
$x_3 = 69,343.50$	Yes	$x_{18} = 64,980.00$	Yes
$x_4 = 69,894.90$	Yes	$x_{19} = 71,932.60$	Yes
$x_5 = 67,621.60$	No	$x_{20} = 72,973.00$	Yes
$x_6 = 75,924.00$	Yes	$x_{21} = 65,120.90$	Yes
$x_7 = 69,092.30$	Yes	$x_{22} = 71,753.00$	Yes
$x_8 = 71,404.40$	Yes	$x_{23} = 74,391.80$	No
$x_9 = 70,957.70$	Yes	$x_{24} = 70,164.20$	No
$x_{10} = 75,109,70$	Yes	$x_{25} = 72,973.60$	No
$x_{11} = 65,922.60$	Yes	$x_{26} = 70,241.30$	No
$x_{12} = 77,268.40$	No	$x_{27} = 72,793.90$	No
$x_{13} = 75,688.80$	Yes	$x_{28} = 70,979.40$	Yes
$x_{14} = 71,564.70$	No	$x_{29} = 75,860.90$	Yes
$x_{15} = 76,188.20$	No	$x_{30} = 77,309.10$	No

14