Chapter 7
Sampling and Sampling Distributions

Review: Inferential Statistics

- Managers frequently make decisions based on sample information
- Parameter
  - defined:
  - μ, σ, p
- Statistic
  - defined:
  - $\overline{x}$, s, $\overline{p}$

Point Estimation

- Using a sample statistic to estimate the value of a population parameter

Parameter	Summary Measure	Statistic
μ	Mean	$\overline{x}$
σ^2	Variance	s^2
σ	Std Deviation	s
p	Proportion (RF)	$\overline{p}$

Simple Random Sampling

- Each possible sample combination has the same probability of being selected
- Two general ways to implement
  - draw from a hat (or coffee can)
  - assign each item a serial number, then use a random number generator
- Replacement?
  - Sampling with replacement vs. sampling without replacement

4

Sampling Distribution Of The Mean (SDOM)

- Use $\overline{X}$ to estimate μ
- The value for $\overline{X}$ will vary, depending on which population items are included in the sample, therefore, it is a random variable
  - if the sample is representative...
  - if the sample is biased ...
- When use $\overline{X}$ to estimate μ, the risk of sampling error is a function of:
  - sample size (n)
  - population dispersion (σ)
- Ch 7 Worksheet 1

5

Sampling Distribution Of The Mean (SDOM)

- "The distribution of sample means"
- Describes the possible values for $\overline{X}$ and their probabilities of occurrence when a sample of size n is selected from a specific population
- Ch 7 worksheet 1
  - Population: {2, 3, 4, 5, 8, 11}
  - μ = 5.5
  - σ = 3.096
  - N = 6
  - possible values of $\overline{X}$ when n=5?

SDOM n=5	
$\overline{X}$	**P($\overline{X}$)**
4.4	1/6
5.0	1/6
5.6	1/6
5.8	1/6
6.0	1/6
6.2	1/6
Total	6/6 = 1.000

6

Expected Value of $\overline{X}$

$$\bar{\bar{x}} = \frac{\sum \bar{x}}{K} \qquad E(\bar{x}) = \bar{\bar{x}} = \mu$$

- Grand Mean (nib)
  - the mean of the sample means
- Some sample means are too high, some are too low, some are right on, but the expected value of the sample mean equals the true population mean, μ
- i.e., $\overline{X}$ is an unbiased estimator of μ

7

Standard Error Of The Mean

Standard Deviation of $\overline{X}$

- The standard deviation of the possible values for the sample mean, $\overline{X}$
- This formula nib but is familiar
- Can only use it when know every value in the population and can list each possible sample and its mean
- The more dispersion among the possible values for $\overline{X}$, the higher the ROSE

$$s_{\bar{x}} = \sqrt{\frac{\sum(\overline{X} - \overline{\overline{X}})^2}{K}}$$

- Summary: The standard error of the mean measures the dispersion among the possible values for $\bar{x}$ and indicates the ROSE when using $\bar{x}$ to estimate μ

8

Standard Error Of The Mean

Standard Deviation of $\overline{X}$

- This formula doesn't require that you know every member of the population and each possible sample mean... although it does require σ
- The second term is called the finite population correction factor. Its value will be ≤ 1 so it reduces the standard error of the mean.
- FCF is waived when
  - (1) infinite population, or
  - (2) sample with replacement, or
  - (3) sample w/o replacement but n < .05N

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

Computational formula

- Ch 7 worksheet 1
  - Population: {2, 3, 4, 5, 8, 11}
  - μ = 5.5 , σ = 3.096, N = 6

n	std error of the mean
2	
5	0.6191
6	

9

The Impact of Sample Size

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

Computational formula

- As n ↑, standard error of the mean ↓
- As σ ↑, standard error of the mean ↑

Sampling From a Normally Distributed Population

- When sampling from a Normally distributed population with (μ, σ), the SDOM has 3 important properties:
  - the distribution of sample means will itself be Normally distributed
  - $E(\bar{x}) = \mu$
  - $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$
- Ch 7, worksheet 2 problem 6

Sampling From a Non-Normally Distributed Popn

- Figure 7.6, page 289
- Central Limit Theorem
  - As n ↑ $\geq$ 30 (n$\geq$50 if population is skewed), the SDOM will have the same 3 properties:
    - the distribution of sample means will itself be Normally distributed
    - $E(\bar{x}) = \mu$
    - $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$
- Ch 7 worksheet 3: Taxi Driver, Water Consumption

Summary: Sampling Distribution Of The Mean

- Use $\bar{x}$ to estimate μ
- $\bar{x}$ is a random variable that can be described by a probability distribution called the sampling distribution of the mean
- The dispersion of SDOM indicates the ROSE and is called the standard error of the mean
- When sampling from a Normally distributed population with σ known **or** from non-Normally distributed population with σ known and n $\geq$ 30, SDOM has 3 properties:
  - the distribution of sample means will itself be Normally distributed
  - $E(\bar{x}) = \mu$
  - $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$

13

Sampling Distribution of the Proportion

- Use $\bar{p}$ to estimate p
- The value for $\bar{p}$ will depend on which items are sampled and how representative the sample is
- Sampling Distribution of the Proportion (SDOP)
  - "the distribution of sample proportions"
  - describes the possible values for $\bar{p}$ and their probabilities of occurrence
- Follows a BINOmial distribution but if both:
  - np $\geq$ 5 **and**
  - n(1-p) $\geq$ 5 then SDOP is approximately Normally distributed

$$E(\bar{p}) = \frac{\sum \bar{p}}{k} = p$$

14

Standard Error of the Proportion

$$\sigma_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}}\sqrt{\frac{N-n}{N-1}}$$

- Indicates the dispersion among the possible values for $\bar{p}$
- The more spread among the possible values for $\bar{p}$, the higher the ROSE
- Ch 7 worksheet 3: MegaCorp, Financial Aid

15

Summary: Sampling Distribution Of The Proportion

- Use $\bar{p}$ to estimate p
- $\bar{p}$ is a random variable that can be described by a probability distribution called the sampling distribution of the proportion
- The dispersion of SDOP indicates the ROSE and is called the standard error of the proportion
- SDOP has 3 important properties:
  - it will be approx Normally distributed if np$\geq$5 **and** n(1-p)$\geq$5
  - $E(\bar{p}) = \dfrac{\sum \bar{p}}{k} = p$
  - $\sigma_{\bar{p}} = \sqrt{\dfrac{p(1-p)}{n}} \sqrt{\dfrac{N-n}{N-1}}$

16

Major Probability Sampling Methods

- Various processes of selecting a sample where each item in population has a chance of being selected
- Probability theory allows us to measure the risk of sampling error (ROSE)

4 major types of probability samples:

- **a) Simple Random**
  - each possible sample combination has the same probability of being selected as any other possible sample combination
- b) Systematic
  - select elements based on a uniform interval of time, space, order
  - caution: selection pattern may coincide with the factor under study!

17

Major Probability Sampling Methods

4 major types of probability samples: (continued)

- c) Stratified
  - divide population into homogenous strata subgroups, take a simple random sample from each; then combine
  - removes a source of bias by taking strata into account and using them to guide proper composition of sample
- d) Cluster
  - divide population into heterogeneous clusters, randomly select entire clusters & measure every member from each selected cluster
  - makes sampling more convenient

18

Experimental Design

- Polls/Studies are carefully planned
- Major Steps:
- 1) a claim is made or an idea surfaces to be tested
- 2) response variable(s) determined
- 3) determine required sample size (Ch 8)
- 4) conduct the study, holding other factors constant
  - matched samples or repeated measures (Ch 10)
- 5) analyze the sample data
- 6) draw conclusions/make a decision

19