Chapter 8
Interval Estimation

Estimation Concepts

- Usually can't take a census, so we must make decisions based on sample data $\bar{x} \to \mu \quad s \to \sigma \quad \bar{p} \to p$
- It imperative that we take the risk of sampling error into account when we interpret sample results
- Point Estimate
  - use a single value to estimate a parameter's value
- Interval Estimate
  - use a range of values to estimate a parameter's value
  - begin with a sample statistic, then set up an interval around it to acknowledge the risk of sampling error
  - eg: poll indicates 65% support proposition 101 with ? 3% point margin of error provides the interval 62% - 68%

2

General Form of a Confidence Interval

$$\text{CI} = \text{sample statistic} \pm \boxed{\text{margin of error}}$$

$$\text{CI} = \text{sample statistic} \pm (\text{interval coefficient})(\text{standard error})$$

- Eg: 65% of those polled support proposition 101 with ? 3% margin of error

.62 ← .65 → .68

3

Confidence Level

- aka Confidence Coefficient
- Commonly used: 90%, 95%, 99%
- Interpretation #1
  - indicates the degree of confidence that the interval contains the true value of the parameter being estimated
- Interpretation #2
  - indicates the percentage of intervals which would include the true value of the parameter being estimated
  - some samples are more biased than this confidence interval process adjusts for
- Level of Significance
  - α = 1- confidence coefficient
  - indicates the probability that the interval does not contain the true value of the parameter being estimated

4

Confidence Level (continued)

- Tradeoff: as the confidence level is increased, the interval widens and less informative
- Precision (nib)
  - the width of the confidence interval (UCL-LCL)
  - other things being equal, we prefer tight intervals since they're more informative
- Car repair estimate
  - could cost between $100-150
  - will probably cost between $25-500
  - will certainly cost between $0-1,000
- Poll example
  - .65 ± .03 margin of error interval .62 – .68
  - precision = .06 (i.e., 6% total width)

5

Review Chapter 7

Using $\bar{X}$ to Estimate μ

- Sampling Distribution of the Mean (SDOM)
  - the distribution of possible sample means
  - follows z distribution when:
    - sampled population is Normally distributed with σ known

    or
    - sampled population is not Normally distributed but $n \geq 30$ (CLT)
- Standard Error of the Mean
  - the dispersion of the possible values of the sample mean
  - indicates ROSE

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

6

Estimating a Population Mean with σ "Known"

- In reality, σ will never be known
  - if you have historical data you can estimate s reliably and independently of the sample you use to estimate the mean

$$CI\ for\ \mu = \bar{x} \pm z\sigma_{\bar{x}} \quad where\ \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

margin of error

- This process can be used when:
  - you have historical data to estimate σ AND
  - either (any shaped population, provided $n \geq 30$)
  - or ($n \geq 15$ and symmetric population)
- ADOT assignment

pg316

7

Estimating a Population Mean with σ **Un**known

- When σ is unknown, we must use sample data to estimate it with s
- But now we're using the same sample for 2 purposes:
  - to calculate $\bar{x}$ in order to estimate **μ**
  - to calculate **s** in order to estimate **σ**

$$CI\ for\ \mu = \bar{x} \pm t_{n-1}\, s_{\bar{x}} \quad where\ s_{\bar{x}} = \frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

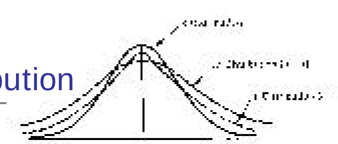
margin of error

- This process can be used when:
  - you lack historical data to estimate σ AND
  - either (Normal population, regardless of n)
  - or (symmetric population, provided $n \geq 15$)
  - or (skewed population, provided $n \geq 50$)

pg323+

8

t-Distribution



- Looks like a flattened z distribution
  - has more area out in its tails
  - wider intervals are produced with t than with z
- t-distribution's shape depends on the degrees of freedom associated with the sample
  - df=n-1 here because we are using the same sample for two purposes:
    - $s \rightarrow \sigma$
    - $\bar{x} \rightarrow \mu$
- OKAYDATA assignment

9

Summary: Constructing Confidence Intervals For μ

- Figure 8.9 page 325

$$\text{CI for } \mu = \overline{X} \pm \begin{pmatrix} t_{n-1} \\ z \end{pmatrix} \begin{pmatrix} \sigma_{\overline{X}} \\ s_{\overline{X}} \end{pmatrix}$$

10

Controlling The M.O.E.

- Margin of Error
  - when using sample mean to estimate μ:

$$\text{M.O.E.} = z \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

- Reduce M.O.E. by:
  - ↓ confidence level
  - ↑ sample size

11

Determining the Sample Size Required To Estimate μ

margin of error

$$\text{CI for } \mu = \overline{x} \pm z\frac{\sigma}{\sqrt{n}} \longrightarrow n_{min} = \frac{z^2 \sigma^2}{E^2}$$

- always round ↑ (& n $\geq$ 30 unless you believe the population is Normally distributed)
- to determine required sample size, specify:
  - confidence level
  - s (estimate using s historical data or a pilot sample)
  - desired margin of error (a.k.a. tolerable error)
- ADOT assignment revisited

12

Review Chapter 7

Using $\overline{p}$ to Estimate p

- Sampling Distribution of the Proportion (SDOP)
  - the distribution of possible values for $\overline{p}$
  - follows a BINOmial distribution, but can use z when:
    - np $\geq$ 5 **and** n(1-p) $\geq$ 5
- Standard Error of the Proportion
  - measures ROSE when use $\overline{p}$ to estimate p

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

  - but you'll never know p... can only estimate it with $\overline{p}$, thus can only estimate the standard error of the proportion

$$s_{\overline{p}} = \sqrt{\frac{\overline{p}(1-\overline{p})}{n}}\sqrt{\frac{N-n}{N-1}}$$

13

Interval Estimation of a Population Proportion (p)

$$\text{CI for } p = \overline{p} \pm z\sqrt{\frac{\overline{p}(1-\overline{p})}{n}}\sqrt{\frac{N-n}{N-1}}$$

iff: np $\geq$ 5 and n(1-p) $\geq$ 5

- Highway Funds assignment

14

Determining The Sample Size To Estimate p

$$\text{CI for } p = \overline{p} \pm z\sqrt{\frac{\overline{p}(1-\overline{p})}{n}} \rightarrow n_{min} = \frac{z^2\, p^*(1-p^*)}{E^2}$$

margin of error

- always round ↑
- to determine necessary sample size, specify:
  - confidence level
  - p (estimate from pilot sample or use .50 to be conservative)
  - desired margin of error (a.k.a. tolerable error)
- Highway Funds, revisited
- Arizona Republic Poll Methodology

15