Point Estimation, Large-Sample C.I.s for a Population Mean

Chapter 7: Estimation and Statistical Intervals
7.1 Point Estimation

• In Statistics, in many cases we are interested in a certain population and its parameter(s)

• To get big picture about the population, we measure statistics from random samples

• Examples of parameters vs. statistics: $\mu$ vs. sample mean, $\sigma$ vs. sample s.d., etc
Point Estimators

• Sample statistics, when used to estimate population parameters, are called “point estimators”

  – The statistic $\bar{X}$ is a point estimate for $\mu$, etc.

  – Other examples?
Unbiased (ideally)

• Estimates can be *unbiased*
  – From chapter 5, we know that the mean \( \mu \) is \( \overline{X} \).
  – If the mean of an estimate is the population parameter it estimates, we call that estimate “unbiased”.
  – So clearly \( \overline{X} \) is an unbiased estimate of \( \mu \).
  – See Figure 7.1 on Page 292, in the textbook

• \( \hat{P} \) is also an unbiased estimate of ____?
• Caution: not all “good” estimators are unbiased, but in most cases it’s preferred when statistics are unbiased.
Consistent Estimators

• Again consider $\bar{X}$, recall that the standard deviation of $\bar{X}$ is $\sigma / \sqrt{n}$.

• What happens to the standard deviation as the sample size:
  – increases?
  – goes to infinity?

• If an estimator converges to the population parameter, with probability 1, when the sample size increases, we call the estimator a consistent estimator.

• $\bar{X}$ is a consistent estimator for $\mu$
7.2 Large Sample Confidence Intervals for a Population Mean

• **Example 1:**
  – Suppose we observe 41 plots of corn with yields (in bushels), randomly selected
  – Sample Mean, $\bar{X} = 123.8$
  – Sample Standard Deviation $= 12.3$
  – What can be said about the (population) mean yield of this variety of corn?
Sampling Distribution

• Assume the yield is $N(\mu, \sigma)$ with unknown $\mu$ and $\sigma$

• Then $\bar{X} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$

• Now while we don’t know $\sigma$, we can replace it with the sample standard deviation, $s$.

• It turns out that when $n$ is large, replacement of $\sigma$ with $s$ does not change much (for now)

• So,

$$s_{\bar{X}} = \frac{s}{\sqrt{n}} = \frac{12.3}{\sqrt{41}} = 1.92$$

Lecture 11
2σ rule

- 68-95-99.7% rule: 95% of time sample mean is approximately within 2 standard deviations of population mean
  
  $- 2 \times 1.92 = 3.84 \text{ from } \mu$

- Thus, 95% of time:
  $\mu - 3.84 < \bar{X} < \mu + 3.84$
Figure 7.3 Capturing a central $z$ curve area of .95
Put differently: 95% of time

\[ \bar{X} - 3.84 < \mu < \bar{X} + 3.84 \]

• In the long-run (with large number of repeated samplings), the random interval covers the unknown (but nonrandom) population parameter \( \mu \) 95% of time.
• Our confidence is 95%.
• We need to be extremely careful when observing this result.
Example 1 (cont):

\[ \bar{X} \pm 3.84 \]

(119.96, 127.64)

- This *particular* confidence interval may contain \( \mu \) or not...
- However, such a systematic **method** gives intervals covering the population mean \( \mu \) in 95% of cases.
- Each interval is NOT 95% correct. Each interval is 100% correct or 100% wrong.
  - It’s the **method** that is correct 95% of the time
Figure 7.5 95% confidence intervals for $\mu$ from 100 different samples (* identifies an interval that does not include $\mu$)
Confidence Intervals (CIs):

• Typically: estimate ± margin of error

• Always use an interval of the form (a, b)

• Confidence level (C) gives the probability that such interval(s) will cover the true value of the parameter.
  – It does not give us the probability that our parameter is inside the interval.
  – In Example 1: C = 0.95, what Z gives us the middle 95%? (Look up on table)
    Z-Critical for middle 95% = 1.96
  – What about for other confidence levels?
    • 90%? 99%?
    • 1.645 and 2.575, respectively.
A large-sample Confidence Interval:

• Data: SRS of n observations (large sample)
• Assumption: population distribution is \( N(\mu, \sigma) \) with unknown \( \mu \) and \( \sigma \)
• General formula:

\[
\bar{X} \pm (z \text{ critical value}) \frac{s}{\sqrt{n}}
\]
After Class...

• Read Sec 7.1 and 7.2, understand the meaning of “confidence level”.

• Review Ch.1, 2 and 5. Make your own Cheat-Sheet (one page, handwritten)
  – Practice Test
  – Review notes, hw and conceptual Qs in Lab