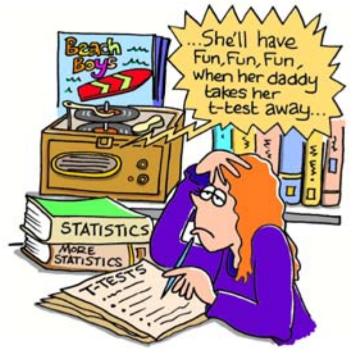
Confidence Interval



Today's Class

- Point Estimate
- Confidence Interval
  - Distribution: Normal vs. Not normal
  - Variance: known vs. unknown
  - Sample size: small vs. large

Estimating a Population Parameter

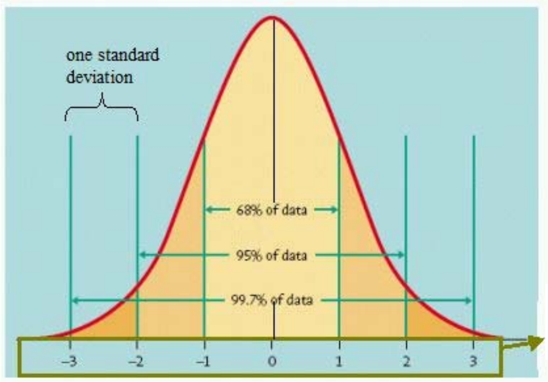
- What is the population mean?
  - Don't know μ? Estimate it.
  - How?
    - Take a sample (n=?)
    - Use $\bar{X}$ to estimate μ
- What is a point estimate?
  - A point estimate is a sample statistic used to estimate the corresponding population parameter

What is Point Estimate of the True Population Mean?

- Use the CLT to know
  - The sample mean ≈ the population mean
  - 68% of all possible sample means drawn from samples you took should be within one standard error of the mean
  - The Standard error = $\frac{\sigma_X}{\sqrt{n}}$
  - So take a large sample and you should have a sample mean very close to μ

Normal Curve, revisit

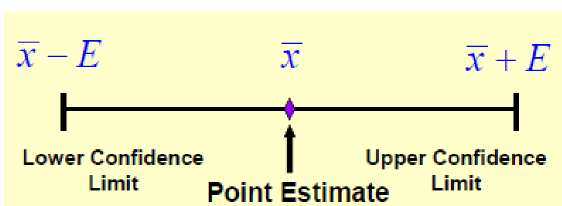
- Approximate percentage of area within given standard deviations



$$Z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}}$$

Confidence Intervals

- Developed from sample data
- If all possible intervals of a given width were constructed, a percentage of these intervals, known as the ***confidence level***, would include the true population parameter



Confidence Intervals Example

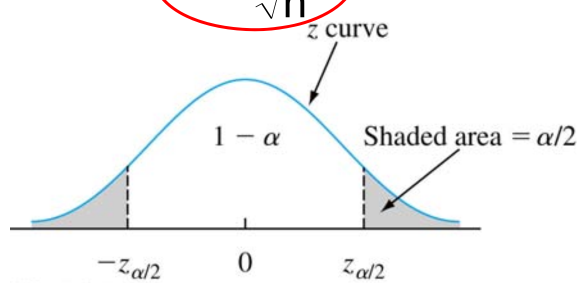
- The internal pressure strength of glass bottles used to package a carbonated beverage is an important quality characteristic
  - Standard deviation: 10 psi
  - Sample mean: 182 psi
  - Sample size: 25
  - Find the 95% confidence interval

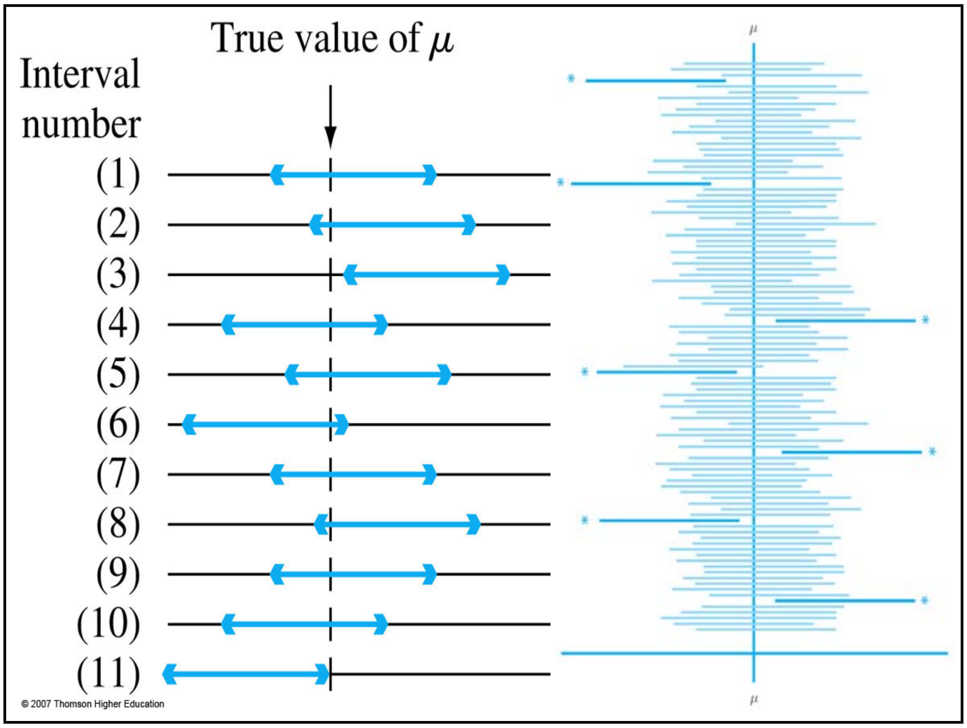
Confidence Intervals: Normal, STD known

- A 100(1-α)% confidence interval for the mean μ of a normal population when the value of σ is known:

$$\bar{x} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Margin of Error (ME)





Interpreting a Confidence Interval

- A correct interpretation of "95% confidence" relied on the long-run relative frequency interpretation of probability
- Suppose we obtain another sample and compute another 95% interval, and so on. In the long run 95% of our computed CIs will contain

Confidence Intervals Example

- The internal pressure strength of glass bottles used to package a carbonated beverage is assumed to be normal
  - Sample size: 25
  - Sample mean: 182 psi
  - Sample standard deviation: 10 psi
  - 95% two-sided confidence interval?

Confidence Interval for Normal, STD Unknown

- A 100(1-a)% two-sided confidence interval for the mean μ of a normal population with $\bar{x}$, the sample mean and s, the sample standard deviation from a random sample of size n:

$$\bar{x} \pm t_{\alpha/2,n-1} \cdot \frac{s}{\sqrt{n}}$$

T-distribution

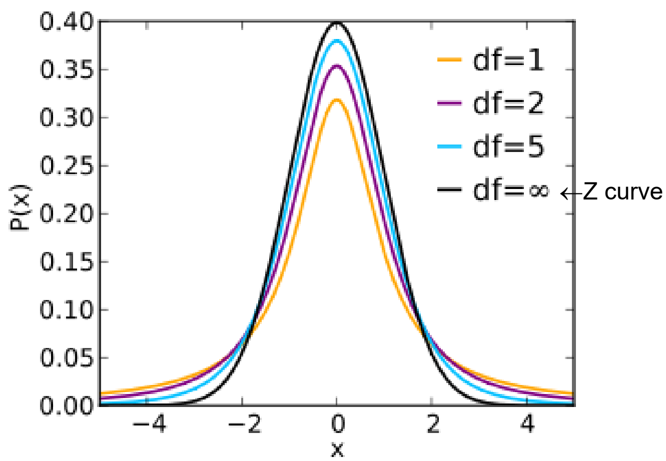
- When $\overline{X}$ is the mean of a random sample of size n from a Normal with mean μ, then

$$T = \frac{\overline{x} - \mu}{S / \sqrt{n}}$$

where S is sample standard deviation

- A t distribution (Appendix A.5) has one parameter, ν = n-1 degrees of freedom

T and Z Curve



Properties of *t* Distributions

- Let t_n denote the t distribution with n df.
  - Each t_n curve is bell-shaped and centered at 0
  - Each t_n curve is more spread out than the standard normal (z) curve
  - As n increases, the spread of the corresponding t_n curve decreases
  - As $n \to \infty$, the sequence of t_n curves approaches the standard normal curve

  (so the z curve is often called the t curve with df = ∞).

Example

- Consider the following sample of fat content (in percentage) of n = 10 randomly selected hot dogs ("Sensory and Mechanical Assessment of the Quality of Frankfurters," *J. of Texture Studies,* 1990: 395–409):

25.2	21.3	22.8	17.0	29.8	21.0
25.5	16.0	20.9	19.5		

  Assuming that these were selected from a normal population distribution, what is a 95% CI for the population mean fat content?

Large Sample CI: Not Necessarily Normal

- If n is sufficiently large (the CLT applies) then a large-sample confidence interval for μ with confidence level approximately 100(1-α)% is:

$$\bar{x} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

  - Generally about 30 observations

Large Sample CI: Not Necessarily Normal, STD Unknown

- If n is sufficiently large (the CLT applies) then a large-sample confidence interval for μ with confidence level approximately 100(1-α)% is:

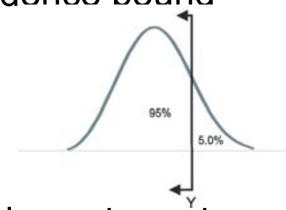
$$\bar{x} \pm z_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

  - Note that it uses the sample STD
  - Generally at least 40 observations are needed

One-Sided Confidence Interval

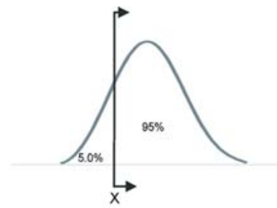
- A large-sample upper confidence bound for μ:

$$\mu < \bar{x} + z_\alpha \cdot \frac{s}{\sqrt{n}}$$



- A large-sample lower confidence bound for μ:

$$\mu > \bar{x} - z_\alpha \cdot \frac{s}{\sqrt{n}}$$



Example One-sided CI

- A sample of 48 shear strength observations gave a sample mean strength of 17.17 N/mm^2 and a sample standard deviation of 3.28 N/mm^2. Find a lower confidence bound for true average shear strength μ with confidence level 95%.