Estimating $\mu$ With Unknown $\sigma$

- This is often true in practice.
- When the sample is large and $\sigma$ is unknown, the sampling distribution is approximately normal regardless of the underlying distribution.
- When the sample is small and $\sigma$ is unknown, we must make an assumption about the form of the underlying distribution to obtain a valid CI procedure.
- A reasonable assumption in many cases is that the underlying distribution is normal.
A Large-Sample CI for $\mu$ (Unknown $\sigma$)

When $n$ is large, the quantity

$$\frac{\bar{X} - \mu}{S/\sqrt{n}}$$

has an approximate standard normal distribution. Consequently,

$$\bar{X} - z_{\alpha/2} \frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + z_{\alpha/2} \frac{S}{\sqrt{n}}$$  \hspace{1cm} (8-11)

is a large sample confidence interval for $\mu$, with confidence level of approximately $100(1 - \alpha)\%$.

Cl on $\mu$ of a Normal Distribution (Unknown $\sigma$)

- Let $\bar{X}$ and $S^2$ be the sample mean and variance of a random sample from a normally distributed population with unknown $\mu$ and $\sigma$. The random variable

$$T = \frac{\bar{X} - \mu}{S / \sqrt{n}}$$

has a Student-$t$ distribution with $n - 1$ degrees of freedom.
**Student-\(t\) Distribution**

\[
f_t(x) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{\pi k} \, \Gamma\left(\frac{k}{2}\right)} \left(\frac{x^2}{k} + 1\right)^{-\frac{k+1}{2}}, \quad -\infty < x < \infty \text{ and } k = \text{df}
\]

\[
E(X) = 0
\]

\[
\text{Var}(X) = \frac{k}{k-2}, \quad k > 2
\]

---

**The \(t\) Confidence Interval on \(\mu\)**

If \(\bar{x}\) and \(s\) are the mean and standard deviation of a random sample from a normal distribution with unknown variance \(\sigma^2\), a \(100(1 - \alpha)\)% confidence interval on \(\mu\) is given by

\[
\bar{x} - t_{\alpha/2, n-1}s/\sqrt{n} \leq \mu \leq \bar{x} + t_{\alpha/2, n-1}s/\sqrt{n}
\]

(8-16)

where \(t_{\alpha/2, n-1}\) is the upper \(100\alpha/2\) percentage point of the \(t\) distribution with \(n - 1\) degrees of freedom.
Which (if any) distribution to use?

Procedure for the interval estimation of $\mu$ with $\sigma$ unknown
A manager of a paint store, wants to estimate the mean amount of a product sold per day. Twenty business days are monitored, and an average of 32 gallons is sold daily. The sample standard deviation is 12 gallons. Calculate the confidence limits at the 95% confidence level.
A random sample has been taken from a normal population with the following statistics:

<table>
<thead>
<tr>
<th>N</th>
<th>Mean</th>
<th>SEmean</th>
<th>Stdev</th>
<th>Variance</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>?</td>
<td>0.507</td>
<td>?</td>
<td>?</td>
<td>251.848</td>
</tr>
</tbody>
</table>

(a) Find the missing quantities (?)
(b) Construct a 95% CI on the population mean
Example

The compressive strength of concrete is being tested by a civil engineer. Twelve specimens are tested and the following data are obtained:

2216  2237  2249  2204  2225  2301
2281  2263  2318  2255  2275  2295

(a) Check the assumption that compressive strength is normally distributed.
(b) Construct a 95% two-sided confidence interval on the mean strength.
Estimating the Population Variance

- Variance shows the extent of the spread or scatter in a data set
- It is desirable to know such variation so that steps can be taken to control it
- Tire manufacturer wants to be sure that tires produced are of consistent mileage quality
- A drug company must focus on the potency of the tablets so that some are not unduly weak while others do not produce overdoses

Estimating the Population Variance

- Let $S^2$ be the sample variance of a random sample taken from a normally distributed population, the sampling distribution of the sample variance follows a chi-square distribution, i.e., the RV
  \[ X^2 = \frac{(n - 1)S^2}{\sigma^2} \]
  has a chi-square ($\chi^2$) distribution with $n-1$ degrees of freedom.
Chi-Square Distribution

\[ f_x(x) = \frac{1}{k} \frac{k^{k/2} x^{k-1} e^{-x/2}}{2^{k/2} \Gamma(k/2)} \], \quad x > 0 \text{ and } k = df

\[ E\{X\} = k \]
\[ Var\{X\} = 2k \]
CI on the Variance of a Normal Population

If $s^2$ is the sample variance from a random sample of $n$ observations from a normal distribution with unknown variance $\sigma^2$, then a $100(1 - \alpha)\%$ confidence interval on $\sigma^2$ is

\[
\frac{(n - 1)s^2}{\chi^2_{\alpha/2,n-1}} \leq \sigma^2 \leq \frac{(n - 1)s^2}{\chi^2_{1 - \alpha/2,n-1}}
\]  

(8-19)

where $\chi^2_{\alpha/2,n-1}$ and $\chi^2_{1 - \alpha/2,n-1}$ are the upper and lower $100\alpha/2$ percentage points of the chi-square distribution with $n - 1$ degrees of freedom, respectively. A confidence interval for $\sigma$ has lower and upper limits that are the square roots of the corresponding limits in Equation 8-19.

Example: A Pressing Problem

The strength and conditioning coach of the Citadel football team described the outcomes of a weight-lifting and fitness program he designed. As part of the program’s evaluation, he had each player do a one-repetition, maximum-weight bench press. The weights pressed by the linebackers are:

340, 380, 305, 335, 375, 400, 305, 385, and 315.

Construct a $90\%$ confidence interval for the standard deviation in the maximum weights pressed by the population of linebackers who go through this program.