1. A survey of first-time home buyers provided a sample mean of annual household income of $50,000 (CNBC.com, July 11, 2000). Assume the survey used a sample of 400 first-time home buyers and that the sample standard deviation of income was $20,500.

   a. Using the 95% confidence coefficient, what is the margin of error for this study?
   b. What is the 95% confidence interval for the population mean annual household income for first-time buyers?
2. The Ford Taurus is listed as having a highway fuel efficiency average of 30 miles per gallon (MPG) (source: 1995 Motor Trend New Car Buyer’s Guide). A consumer interest group conducts automobile mileage tests seeking statistical evidence to show that automobile manufacturers overstate the MPG ratings for particular models. In the case of the Ford Taurus, hypotheses for the test would be stated as $H_0$: $\mu \geq 30$ and $H_a$: $\mu < 30$. In a sample of 50 mileage tests with the Ford Taurus, the consumer interest group obtained a sample mean highway mileage rating of 29.5 MPG and a sample standard deviation of 1.8 MPG. What conclusion should be drawn from the sample results? Use the 1% level of significance.
3. Stout Electrical Company operates a fleet of trucks that provide electrical service to the construction industry. Monthly mean maintenance cost has been $75 per truck. A random sample of 40 trucks provided a sample mean maintenance cost of $82.50 per month with a population standard deviation of $30. Managers want a test to determine whether the mean monthly maintenance cost has increased. Using the 5% level of significance, test the following null hypothesis: $H_0: \mu \leq 75$. 
4. Safegate Foods is redesigning the checkout lanes in its super markets throughout the country. Two designs have been suggested. Tests on customer checkout times have been conducted at two stores where the two new systems have been installed. A summary of the sample data is as follows:

<table>
<thead>
<tr>
<th>System</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>120</td>
<td>100</td>
</tr>
<tr>
<td>Average</td>
<td>4.1</td>
<td>3.3</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>2.2</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Assume that the population standard deviations, which are unknown, are believed to be the same.

a. What is the pooled variance for this question?

b. Test whether there is a difference between the mean checkout times of the two systems. Use the 5% level of significance. Based on the results of your tests, which system is preferred?
5. Home mortgage rates for 30-year fixed-rate loans vary throughout the country. During the summer of 2000, data available from various parts of the country suggested that the standard deviation of the interest rates was 0.096 in 2000 (Wall Street Journal, Sept 8th, 2000). The corresponding variance for 2000 is $0.096^2 = 0.009216$. Consider a follow-up study conducted during the summer of 2001. The interest rates for 30-year fixed rate loans at a sample of 20 lending institutions provided a sample standard deviation of 0.114 (variance = $0.114^2 = 0.012996$). Conduct a hypothesis test of $H_0: \sigma^2 = 0.009216$ to see whether the sample data indicate that the variance in interest rates has changed from 2000 to 2001. Use the 5% level of significance. What is your conclusion?
6. Most individuals are aware that the average annual repair cost for a car depends on the age of the car. A researcher is interested in finding whether the variance of the annual repair costs also increases with the age of the car. A sample of 25 cars 4 years old showed a sample standard deviation of $170 and a sample of 25 cars 2 years old showed a sample standard deviation for annual repair costs of $100. Perform a hypothesis test on whether the variance of 4 year old cars is greater than the variance of 2 year old cars. The null hypothesis for this study is $H_0: \sigma_{\text{old}} \leq \sigma_{\text{new}}$. Use the 1% level of significance. What is your conclusion?
Detailed Answers:

1. A survey of first-time home buyers provided a sample mean of annual household income of $50,000 (CNBC.com, July 11, 2000). Assume the survey used a sample of 400 first-time home buyers and that the sample standard deviation of income was $20,500.

   a. Using the 95% confidence coefficient, what is the margin of error for this study?

   **Answer:** Since we do not know $\sigma$, we use the t-distribution for this problem with 1-\(\alpha = 0.95\).

   Since \(n = 400\), use d.f. = \(\infty\) for this problem: \(t_{0.025} = 1.96\).

   The estimated standard error is \(s_{\bar{x}} = \frac{s}{\sqrt{n}} = \frac{20,500}{\sqrt{400}} = 1,025\).

   Thus, MOE = 1.96*1,025 = 2,009

   b. What is the 95% confidence interval for the population mean annual household income for first-time buyers?

   Upper limit: \(\bar{x} + MOE = 50,000 + 2,009 = 52,009\)

   Lower limit: \(\bar{x} - MOE = 50,000 - 2,009 = 47,991\)
2. The Ford Taurus is listed as having a highway fuel efficiency average of 30 miles per gallon (MPG) (source: 1995 Motor Trend New Car Buyer’s Guide). A consumer interest group conducts automobile mileage tests seeking statistical evidence to show that automobile manufacturers overstate the MPG ratings for particular models. In the case of the Ford Taurus, hypotheses for the test would be stated as $H_0$: $\mu \geq 30$ and $H_a$: $\mu < 30$. In a sample of 50 mileage tests with the Ford Taurus, the consumer interest group obtained a sample mean highway mileage rating of 29.5 MPG and a sample standard deviation of 1.8 MPG. What conclusion should be drawn from the sample results? Use the 1% level of significance.

**Answer:** This is a left-tailed test. Since you don’t know the population standard deviation, use the t-distribution in your test.

In this problem, d.f. = $n - 1 = 49$. You could use either d.f. = 40 or d.f. = 60 to answer this problem. I will choose 40 since it is slightly closer to our actual d.f.

The critical value of t at the 1% level of significance is $t_c = -2.423$.

The value of our test statistic is

$$t = \frac{x - \mu}{s/\sqrt{n}} = \frac{29.5 - 30}{1.8/\sqrt{50}} = -1.96$$

Since $-1.96 > -2.423$ we cannot reject the null hypothesis. Our conclusion is that Ford’s information is accurate.
3. Stout Electrical Company operates a fleet of trucks that provide electrical service to the construction industry. Monthly mean maintenance cost has been $75 per truck. A random sample of 40 trucks provided a sample mean maintenance cost of $82.50 per month with a population standard deviation of $30. Managers want a test to determine whether the mean monthly maintenance cost has increased. Using the 5% level of significance, test the following null hypothesis: $H_0: \mu \leq 75$.

**Answer**: This is a right-tailed test. Since you know the population standard deviation, you use the standard-normal distribution.

The critical value of $z$ at the 5% level of significance is $z_c = 1.645$

The value of our test statistic is $z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{82.50 - 75}{30/\sqrt{40}} = 1.58$

Since $1.58 < 1.645$, we cannot reject the null hypothesis.

Note that the p-value for this problem is $0.5 - 0.4429 = 0.0571$. Since $p-value > \alpha$ we cannot reject the null hypothesis by this criterion. This is the same conclusion we reached by comparing our test statistic to our critical value. We will always reach the same conclusion if we compare a test statistic to a critical value or if we compare a p-value to the level of significance.
4. Safegate Foods is redesigning the checkout lanes in its super markets throughout the country. Two designs have been suggested. Tests on customer checkout times have been conducted at two stores where the two new systems have been installed. A summary of the sample data is as follows:

<table>
<thead>
<tr>
<th>System A</th>
<th>System B</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>120</td>
</tr>
<tr>
<td>Average</td>
<td>4.1</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>2.2</td>
</tr>
</tbody>
</table>

Assume that the population standard deviations, which are unknown, are believed to be the same.

a. What is the pooled variance for this question?

Answer: 

\[
\begin{align*}
\hat{s}^2 &= \frac{(n_1-1) * s_1^2 + (n_2-1) * s_2^2}{n_1 + n_2 - 2} \\
&= \frac{119 * 2.2^2 + 99 * 1.5^2}{120 + 100 - 2} = 3.66
\end{align*}
\]

b. Test whether there is a difference between the mean checkout times of the two systems. Use the 5% level of significance. Based on the results of your tests, which system is preferred?

Answer: from part a, the estimated standard deviation is \( s = \sqrt{3.66} = 1.91 \)

The null hypothesis is \( H_o : \mu_1 - \mu_2 = 0 \), giving us a two-tailed test. Since we don’t know the population standard deviations, we must use the t-distribution.

Critical values: with d.f. = 218, use d.f. = \( \infty \) in the table. \( t_c = \pm 1.96 \)

Test statistic: 

\[
\begin{align*}
t &= \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1^0 - \mu_2^0)}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \\
&= \frac{(4.1 - 3.3) - 0}{1.91 \sqrt{\frac{1}{120} + \frac{1}{100}}} \\
&= \frac{0.8}{0.259} = 3.09
\end{align*}
\]

Since \( t > t_c \), we reject the null hypothesis.

Based on this result, assuming the costs of installing the new system are sufficiently small, then I would prefer the new system.
5. Home mortgage rates for 30-year fixed-rate loans vary throughout the country. During the summer of 2000, data available from various parts of the country suggested that the standard deviation of the interest rates was 0.096 in 2000 (Wall Street Journal, Sept 8th, 2000). The corresponding variance for 2000 is $0.096^2 = 0.009216$. Consider a follow-up study conducted during the summer of 2001. The interest rates for 30-year fixed rate loans at a sample of 20 lending institutions provided a sample standard deviation of 0.114 (variance = $0.114^2 = 0.012996$).

Conduct a hypothesis test of $H_0: \sigma^2 = 0.009216$ to see whether the sample data indicate that the variance in interest rates has changed from 2000 to 2001. Use the 5% level of significance. What is your conclusion?

**Answer:** We have a two-tailed test in this case. Since we are examining a population variance, we use the chi-squared distribution.

Critical values of the chi-squared distribution for this problem: d.f. = 20-1 = 19

$\chi^2_{0.975} = 8.90655$ defines the left - tail rejection region

$\chi^2_{0.025} = 32.8523$ defines the right - tail rejection region

Test statistic: $\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{19 \times 0.012996}{0.009216} = 26.79$

Since $8.90655 < 26.79 < 32.8523$, our test statistic does not lie in either rejection region, so we do not reject the null. Our conclusion is that the variance in interest rates has not changed.
6. Most individuals are aware that the average annual repair cost for a car depends on the age of the car. A researcher is interested in finding whether the variance of the annual repair costs also increases with the age of the car. A sample of 25 cars 4 years old showed a sample standard deviation of $170 and a sample of 25 cars 2 years old showed a sample standard deviation for annual repair costs of $100. Perform a hypothesis test on whether the variance of 4 year old cars is greater than the variance of 2 year old cars. The null hypothesis for this study is $H_0: \sigma_{\text{old}} \leq \sigma_{\text{new}}$. Use the 1% level of significance. What is your conclusion?

**Answer:** We have a right-tailed test in this problem. Use the F-distribution

Critical value: $F_{24,24} = 2.66$

Test statistic: $F = \frac{170^2}{100^2} = 2.89$

Since $2.89 > 2.66$ (our critical value), we reject the null hypothesis and conclude that the variance of repair costs for 4 year-old cars is larger than that for 2 year-old cars.