Hypothesis Testing

We will distinguish the following two types of hypotheses:

- **Research Hypothesis**: the conjecture or supposition that motivates the research.

- **Statistical Hypothesis**: it is stated in such a way it may be evaluated by appropriate statistical techniques. It is usually formulated based on the research hypothesis.
Example: Suppose we want to test whether the mean serum cholesterol level $\mu$ of the subpopulation of hypertensive smokers is equal to the mean $\mu_0 = 211$ of the general population of 20- to 74-year-old males. We state the hypotheses as

$$H_0 : \mu = 211,$$

$$H_A : \mu \neq 211.$$ 

We call $H_0$ the null hypothesis, and $H_A$ the alternative hypothesis. Whether or not we reject the null hypothesis $H_0$ will be based on the data.

The data consists of a random sample of 12 hypertensive smokers. The sample mean $\bar{x} = 217$. Suppose that the standard deviation $\sigma = 46$ is known.
An outline of hypothesis test

1. Statistical hypothesis (formulated based on research hypothesis)
2. Study design
3. Data collection
4. Statistical model assumptions
5. Test statistic: calculation, distribution
6. Decision rule
7. Conclusion and caveats
Test statistic

In many cases, the test statistic takes the form:  
\[
\text{test statistic} = \frac{\text{relevant statistic} - \text{hypothesized parameter}}{\text{standard error of the relevant statistic}}.
\]

In our example, we use  
\[
Z = \frac{\overline{X} - \mu_0}{\sigma \sqrt{n}}.
\]

The observed value of the test statistic:  
\[
z = \frac{\overline{x} - \mu_0}{\sigma \sqrt{n}} = \frac{217 - 211}{46/\sqrt{12}} = 0.45.
\]

What can we say based on this observed value of the test statistic?
Decision rule

<table>
<thead>
<tr>
<th>Possible action</th>
<th>Null Hypothesis $H_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>True</td>
</tr>
<tr>
<td>Do not reject $H_0$</td>
<td>Correction Decision (Specificity)</td>
</tr>
<tr>
<td>Reject $H_0$</td>
<td>Type I error (False positive)</td>
</tr>
</tbody>
</table>

The type I error is measured by

$$\alpha = P(\text{Reject } H_0|H_0 \text{ is true}).$$

We usually call $\alpha$ the type I error. Conventionally, we control the type I error at a certain level, say 0.05, in making the decision.
Decision rule (continued)

The p-value

The p-value for a hypothesis test is the probability of obtaining, using the distribution under the null hypothesis \(H_0\), a value of the test statistic as extreme as or more extreme (in the appropriate direction) than the one actually observed.

In the example, observed value of the test statistic is \(z = 0.45\) and \(Z \sim N(0, 1)\). So the p-value is

\[
p = P(|Z| \geq 0.45) = P(Z \leq -0.45) + P(Z \geq 0.45)
= 2P(Z \geq 0.45) = 2 \times 0.326 = 0.652.
\]

Given a threshold value of the type I error rate, say, \(\alpha = 0.05\), we do not reject the null hypothesis \(H_0\) because the p-value > 0.05.
Equivalence between tests and confidence intervals

In this example, the p-value = 0.652. Thus we do not reject the null hypothesis at the 0.05 significance level.

We can calculate that the 95% confidence interval for the mean $\mu$ is

$$(217 - 1.96 \times \frac{46}{\sqrt{12}}, 217 + 1.96 \times \frac{46}{\sqrt{12}}) = (191, 243).$$

Therefore, the null hypothesis value $\mu_0 = 211$ is in the 95% confidence interval.

In general, if we do not reject the null hypothesis $H_0$ at the $\alpha$ significance level, then the null hypothesis value $\mu_0$ is in the $1 - \alpha$ confidence interval. Conversely, if the null hypothesis value $\mu_0$ is contained in the $(1 - \alpha)$ confidence interval, then we do not reject the $H_0$ at the $\alpha$ significance level.
Conclusion In this example, the data does not provide evidence that rejects the null hypothesis that the mean serum cholesterol level in the subpopulation of hypertensive smokers of 20- to 74-year old males is different from the population of 20- to 74-year old males. However, the sample size ($n = 12$) is small. A larger sample and further investigation are needed to give a more definitive answer to the hypothesis.