Chapter 14

Introduction to Inference
The statement being tested in a statistical test is called the **null hypothesis**.

The test is designed to assess the strength of evidence against the null hypothesis.

Usually the null hypothesis is a statement of “no effect” or “no difference”, or it is a statement of equality.

When performing a hypothesis test, we assume that the null hypothesis is true until we have sufficient evidence against it.
Stating Hypotheses
Alternative Hypothesis, $H_a$

- The statement we are trying to find evidence for is called the alternative hypothesis.
- Usually the alternative hypothesis is a statement of “there is an effect” or “there is a difference”, or it is a statement of inequality.

- The alternative hypothesis should express the hopes or suspicions we bring to the data. It is cheating to first look at the data and then frame $H_a$ to fit what the data show.
Tests for a Population Mean

The four steps in carrying out a significance test:

1. State the null and alternative hypotheses.
2. Calculate the test statistic.
3. Find the $P$-value.
4. State your conclusion in the context of the specific setting of the test.
The Hypotheses for Means

- **Null:** $H_0: \mu = \mu_0$
- **One sided alternatives**
  - $H_a: \mu > \mu_0$
  - $H_a: \mu < \mu_0$
- **Two sided alternative**
  - $H_a: \mu \neq \mu_0$
Test Statistic

Testing the Mean of a Normal Population

Take an SRS of size $n$ from a Normal population with unknown mean $\mu$ and known standard deviation $\sigma$. The test statistic for hypotheses about the mean ($H_0 : \mu = \mu_0$) of a Normal distribution is the standardized version of $\bar{X}$:

$$z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$
Assuming that the null hypothesis is true, the probability that the test statistic would take a value as extreme or more extreme than the value actually observed is called the \textit{P-value} of the test.

The smaller the \textit{P-value}, the stronger the evidence the data provide against the null hypothesis. That is, a small \textit{P-value} indicates a small likelihood of observing the sampled results if the null hypothesis were true.
**P-value for Testing Means**

- **$H_a$: $\mu > \mu_0$**
  - P-value is the probability of getting a value as large or larger than the observed test statistic ($z$) value.

- **$H_a$: $\mu < \mu_0$**
  - P-value is the probability of getting a value as small or smaller than the observed test statistic ($z$) value.

- **$H_a$: $\mu \neq \mu_0$**
  - P-value is *two times* the probability of getting a value as large or larger than the absolute value of the observed test statistic ($z$) value.
Statistical Significance

- If the $P$-value is as small as or smaller than the significance level $\alpha$ (i.e., $P$-value $\leq \alpha$), then we say that the data give results that are statistically significant at level $\alpha$.

- If we choose $\alpha = 0.05$, we are requiring that the data give evidence against $H_0$ so strong that it would occur no more than 5% of the time when $H_0$ is true.

- If we choose $\alpha = 0.01$, we are insisting on stronger evidence against $H_0$, evidence so strong that it would occur only 1% of the time when $H_0$ is true.
z TEST FOR A POPULATION MEAN

Draw an SRS of size n from a Normal population that has unknown mean \( \mu \) and known standard deviation \( \sigma \). To test the null hypothesis that \( \mu \) has a specified value,

\[ H_0: \mu = \mu_0 \]

calculate the \textbf{one-sample z statistic}

\[ z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \]

In terms of a variable \( Z \) having the standard Normal distribution, the \( P \)-value for a test of \( H_0 \) against

\[ H_a: \mu > \mu_0 \text{ is } P(Z \geq z) \]

\[ H_a: \mu < \mu_0 \text{ is } P(Z \leq z) \]

\[ H_a: \mu \neq \mu_0 \text{ is } 2P(Z \geq |z|) \]
Case Study I
Sweetening Colas

Diet colas use artificial sweeteners to avoid sugar. These sweeteners gradually lose their sweetness over time. Trained testers sip the cola and assign a “sweetness score” of 1 to 10. The cola is then retested after some time and the two scores are compared to determine the difference in sweetness after storage. Bigger differences indicate bigger loss of sweetness.
Case Study I
Sweetening Colas

Suppose we know that for any cola, the sweetness loss scores vary from taster to taster according to a Normal distribution with standard deviation $\sigma = 1$.

The mean $\mu$ for all tasters measures loss of sweetness.

The sweetness losses for a new cola, as measured by 10 trained testers, yields an average sweetness loss of $\bar{x} = 1.02$. **Do the data provide sufficient evidence that the new cola lost sweetness in storage?**
Case Study I
Sweetening Colas

1. **Hypotheses:**
   - $H_0: \mu = 0$
   - $H_a: \mu > 0$

2. **Test Statistic:**
   $$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{1.02 - 0}{1 / \sqrt{10}} \approx 3.23$$

3. **P-value:**
   - $P$-value = $P(Z > 3.23) = 1 - 0.9994 = 0.0006$

4. **Conclusion:**
   Since the $P$-value is smaller than $\alpha = 0.01$, there is very strong evidence that the new cola loses sweetness on average during storage at room temperature.
Case Study II
Studying Job Satisfaction

Does the job satisfaction of assembly workers differ when their work is machine-paced rather than self-paced? A matched pairs study was performed on a sample of workers, and each worker’s satisfaction was assessed after working in each setting. The response variable is the difference in satisfaction scores, self-paced minus machine-paced.
Case Study II

Studying Job Satisfaction

Suppose job satisfaction scores follow a Normal distribution with standard deviation $\sigma = 60$. Data from 18 workers gave a sample mean score of 17.
Case Study II
Studying Job Satisfaction

1. **Hypotheses:**
   - $H_0: \mu = 0$
   - $H_a: \mu \neq 0$

2. **Test Statistic:**
   
   $$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{17 - 0}{60 / \sqrt{18}} \approx 1.20$$

3. **P-value:**
   
   $$P\text{-value} = 2P(Z > 1.20) = (2)(1 - 0.8849) = 0.2302$$

4. **Conclusion:**

   Since the $P$-value is larger than $\alpha = 0.10$, there is not sufficient evidence that mean job satisfaction of assembly workers differs when their work is machine-paced rather than self-paced.