Hypothesis Test on $\mu$

Large Sample, $\sigma$ unknown

Ex: A company owns a fleet of cars with mean MPG known to be $\mu = 30$. This company will use a gasoline additive only if the additive increases gasoline MPG. A sample of 36 cars used the additive. The sample mean was calculated to be $\bar{X} = 31.3$ miles with std. dev. $s = 7.0$ miles. Does the additive significantly increase mean MPG?

Use $\alpha = .10$. 

1) \( H_0: \mu = 30 \)

2) \( H_a: \mu > 30 \)

3) \( \alpha = .10 \)

4) \( n = 36 \quad \bar{x} = 31.3 \quad s = 7.0 \)

\[
T = \frac{\bar{x} - \mu_0}{s / \sqrt{n}} = \frac{31.3 - 30}{7 / \sqrt{36}} = \frac{1.3}{1.17} = 1.11
\]

5) The p-value is the probability of getting a value of 1.11 or larger for the test stat. (since \( H_a: \mu > 30 \))

The test stat. follows a \( t \)-dist.

But since \( n \) is large
the CLT the CLT tells us that that this $t$-dist will be nearly identical to a standard normal dist. So we can use the standard normal dist. (Table E ) to find our $p$-value;

$$P(T \geq 1.11) \approx P(Z \geq 1.11)$$

$$= 0.1335$$

6) Reject $H_0$ if $p$-value $\leq \alpha$

Since $0.1335 \neq 0.10$

we do not reject $H_0$
7) At the 10% significance level, we conclude that the gasoline additive does not significantly increase MPG of the entire fleet of cars.

Note: Some books call this a Z-test, since we use a standard normal dist. to find the p-value.

Assumptions:
Random sample of independent observations
Hypothesis Test on $\mu$ - Data from Normal Population, $\sigma$ known

The following procedure is appropriate when $\sigma$ is known (unlikely) and the data comes from a Normal pop. The sample size $n$ may be large or small, i.e., for any sample size $n \geq 2$.

Ex: A CNC machine (computerized numeric control) is set to produce rods 10 cm. long. The manufacturer of the CNC machine states that the standard deviation of the rods will always be 0.75 cm, no
no matter the length of the rods produced.
A production technician believes too many "short" rods are being produced, indicating that the CNC machine needs re-calibration.
A random sample of 9 rods produced a sample mean equal to 9.34 cm. Does the data support the technician's belief? You may assume rod lengths are normally distributed. Use α = 0.05.

1) H₀: \( \mu = 10 \)
2) Hₐ: \( \mu < 10 \)
3) \( \alpha = 0.05 \)

4) Test-statistic

\[
Z = \frac{\bar{X} - \mu_0}{\left( \frac{\sigma}{\sqrt{n}} \right)} \\
= \frac{9.34 - 10}{\left( \frac{0.75}{\sqrt{9}} \right)} \\
= \frac{-0.66}{0.25} \\
= -2.64
\]

5) P-value of Test Statistic

Since \( H_A: \mu < 10 \), we must find the probability of getting a test-statistic value \(-2.64\) or smaller, i.e.,
\[ P(z \leq -2.64) = 0.0041 \]

6) Decision

Reject \( H_0 \) if p-value \( \leq \alpha \)
Since \( 0.0041 \leq 0.05 \)
we reject \( H_0 \)

7) Conclude:

At the 5\% significance level, we conclude the mean lengths of rods produced by the CNC machine is significantly less than 10 cm.

Note: Based on this conclusion, we will schedule the CNC machine for re-calibration.
Assumptions:

1) Random Sample of Independent observations from a Normal dist.

2) \( \sigma \) is known
Hypothesis Test on $\mu$ - Small Sample, $\sigma$ unknown, Data From Normal Population

Ex: Automobile exhaust contains an average of 90 parts per million (ppm) of carbon monoxide. A new pollution control device is placed on ten randomly selected cars. For these 10 cars, mean carbon monoxide emission was 75 ppm with a standard deviation of 20 ppm. Does the new pollution control device significantly reduce carbon monoxide emission? Use $\alpha = 0.05$. You may assume the data comes from a Normal Dist.
1) $H_0: \mu = 90$

2) $H_a: \mu \leq 90$

3) $\alpha = 0.05$

4) Test statistic

$$T = \frac{\bar{X} - \mu_0}{(S/\sqrt{n})}$$

$$= \frac{75 - 90}{(20/\sqrt{10})}$$

$$= \frac{-15}{6.3246}$$

$$= -2.37$$

5) P-value of test statistic

Since $H_a: \mu < 90$, we must find the probability of getting a value of $-2.37$ or smaller
for the test-statistic, i.e.,
\[ P(T \leq -2.37) \]
using a \( t \)-dist with \( df = n-1 = 9 \).
Note that, in this example, we cannot use a standard normal dist. to find this probability—since \( n \) is NOT large.
Looking in the \( t \)-dist. table F along the row for 9 d.f., we try to find -2.37, but notice that all values in the \( t \)-table are positive.
But we can use symmetry of \( t \)-dist to help us. Note that 2.37 is between 2.262 (under the column
labeled 0.025) and 2.398 (under the column labeled 0.02)
So our p-value is somewhere between 0.025 and 0.02

6) Decision
Reject Ho if p-value ≤ α

We know that our p-value is p-value ≤ 0.025, and since 0.025 < 0.05 (our α), then p-value ≤ 0.05.
So we reject Ho.

7) At the 5% significance level, the evidence indicates that the new pollution control device
does significantly reduce carbon monoxide emission.

Note: Most statistical computer software produces exact p-values.

Assumptions - Random sample of independent observations from a Normal Dist.