# Lesson 4.3

## Testing Hypothesis About a Population Parameter

Norberto E. Milla, Jr.

## Learning Outcomes

- Perform a test of hypothesis for the population mean
- Perform a test of hypothesis for the population proportion
- Use R to conduct tests of hypothesis
- Interpret results of test of hypothesis

# Test of Hypothesis for the Population Mean

- Null hypothesis: $H_0 : \mu = \mu_0$, where: $\mu_0$ is the hypothesized value of the parameter $\mu$
- Alternative hypothesis (any one of the following)
  - $H_1 : \mu > \mu_0$
  - $H_1 : \mu < \mu_0$
  - $H_1 : \mu \neq \mu_0$
- Test statistic: t test

$$t = \frac{\overline{y} - \mu_0}{\frac{s}{\sqrt{n}}}$$

where: $\overline{Y}$, $s$, and $n$ are the sample mean, sample standard deviation, and sample size, respectively

# Test of Hypothesis for the Population Mean

**Remarks**

- The $T$ test statistic is useful under the following conditions:
  - the sample size ($n$) is small, usually below 30
  - the population standard deviation, $\sigma$, is unknown
  - the data is obtained using random sample from a normally distributed population
- When the sample size is large ($n \geq 30$) and/or the population standard deviation ($\sigma$) is known, then the $Z$ test statistic must be used
- For large samples, the distribution of the $T$ test statistic converges to the distribution of the $Z$ test statistic, that is, $T$ and $Z$ are asymptotically equivalent for large samples

# Test of Hypothesis for the Population Mean: an example

**Problem**:

The average time to complete the pen-and-paper admission test to a university was recorded to be 1.45 hours. Because of the COVID-19 pandemic, the admission test was administered online and it is believed that the average time to complete the test would be shorter than the established average. To test this claim, the completion time of a random sample of 25 examinees was recorded. The sample yielded a mean of 1.23 hours with standard deviation of 0.35 hours. Use a significance level of 0.05.

**Solution**:

- $H_0 : \mu = 1.45$
- $H_1 : \mu < 1.45$
- $\alpha = 0.05$

# Test of Hypothesis for the Population Mean: an example

- **Test statistic**: t
  - Sample evidence: $\overline{y} = 1.23, s = 0.35, n = 25$

$$
\begin{aligned}
t &= \frac{\overline{y} - \mu_0}{s/\sqrt{n}} \\
&= \frac{1.23 - 1.45}{0.35/\sqrt{25}} \\
&\approx -3.14
\end{aligned}
$$

with p-value equal to 0.0022.

- **Rejection Rule**: Reject $H_0$ if $t < -t_{0.05,24} = -1.71$.
- **Decision**: Reject $H_0$ since $-3.14 < -1.71$.
- **Conclusion**: At the 5% level of significance, the data is sufficient to conclude that the average time to complete the admission test was significantly reduced when it is administered online .

# Test of Hypothesis for the Population Mean: an example

**Computing T test statistic and its p-value using R (based on summary statistics)**

```
BSDA::tsum.test(mean.x = 1.23,
                s.x = 0.35,
                n.x = 25,
                mu = 1.45,
                alternative = "less")
```

```
##
##  One-sample t-Test
##
## data:  Summarized x
## t = -3.1429, df = 24, p-value = 0.002205
## alternative hypothesis: true mean is less than 1.45
## 95 percent confidence interval:
##        NA 1.349762
## sample estimates:
## mean of x
##      1.23
```

# Test of Hypothesis for the Population Mean: another example

In a career guidance seminar for graduating economics students, a representative of a BPO firm claims that their employees earn, on average, more than PhP40,000 per month. To validate this claim a researcher calculated the monthly salary of 25 randomly selected employees of the company and obtained a mean salary of PhP45,715 per month with standard deviation PhP11,213. Does the data support the claim of the BPO representative? Use a 5% level of significance.

# Test of Hypothesis for the Population Proportion

Below are claims about a proportion:

- Based on a sample survey, fewer than 25% of all government employees smoke.
- The proportion of people against COVID-19 vaccination exceeds 15%.
- The effectiveness of the vaccine is less than 90%.
- Less than 20% of college students smoke e-cigarettes.
- It is believed that fewer than 10% of university students do not graduate on time.

# Test of Hypothesis for the Population Proportion

- Null hypothesis: $H_0 : p = p_0$, where: $p_0$ is the hypothesized value of the parameter $p$
- Alternative hypothesis (any one of the following)
  - $H_1 : p > p_0$
  - $H_1 : p < p_0$
  - $H_1 : p \neq p_0$
- Test statistic:
  - Z test
  - Binomial test

# Test of Hypothesis for the Population Proportion

**Remarks**:

- The $Z$ test is used under the following conditions:
  - Data (binary) is based on a random sample from a binomial distribution
  - Sample is large and satisfies $np \geq 5$ and $n(1-p) \geq 5$ [Normal approximation!]

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

where: $\hat{p} = \frac{\text{no. of "favorable outcomes"}}{n}$

- The binomial test is appropriate for small samples

# Test of Hypothesis for the Population Proportion: an example

Suppose 200 accounts were audited and 140 were found to be compliant with standard procedures. The manager of the audit firm claims that more than 60% of accounts are compliant with standard procedures. Is the (sample) data sufficient to conclude that indeed more than 60% of accounts are compliant?

**Solution**:

- $H_0 : p = 0.60$
- $H_1 : p > 0.60$
- $\alpha = 0.05$
- **Test statistic**:
  - Check if Z test is appropriate: We obtained before that $\hat{p} = \frac{140}{200} = 0.70$. Thus, $n\hat{p} = 200 \times 0.70 = 140 > 5$ and $n(1 - \hat{p}) = 200 \times (1 - 0.70) = 60 > 5$. This means that we can use the normal approximation.

# Test of Hypothesis for the Population Proportion: an example

- **Test statistic**:

$$
\begin{aligned}
z &= \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \\
&= \frac{0.70 - 0.60}{\sqrt{\frac{0.60(1-0.60)}{200}}} \\
&\approx 2.89
\end{aligned}
$$

- **Rejection Rule**: Reject $H_0$ if $z > z_{0.05} = 1.645$.
- **Decision**: Reject $H_0$ since $2.89 > 1.645$.
- **Conclusion**: At $\alpha = 0.05$, the data is sufficient to conclude that more than 60% of accounts are compliant.

# Test of Hypothesis for the Population Proportion: an example

A random sample of 1500 Filipinos were asked whether they approve or disapprove of the plan of the Philippine government to re-apply for membership of the International Criminal Court (ICC). Of the 1500 surveyed, 660 respond with "approve". Is the data sufficient to conclude that less than 50% of Filipinos approve the plan to re-apply for ICC membership?