STP 226
ELEMENTARY STATISTICS

CHAPTER 9

HYPOTHESIS TESTS FOR ONE POPULATION MEAN

9.1 The Nature of Hypothesis Testing

Q: Does the mean weight of all pretzels packaged by a particular company differ from the advertised weight of 454 grams (g)?
Q: Does the mean age of all cars in use have increased from the 1995 mean of 8.5 years?

A **hypothesis test** allows us to make a decision of judgment about the value of a parameter (e.g. μ)

Hypothesis – a statement that something is true

A **hypothesis test** involves two hypothesis – **Null hypothesis** and **Alternate (Research**) **hypothesis**.

- **Null hypothesis (H_o)**: A hypothesis to be tested

  $H_0 : \mu = \mu_0$

- **Alternate hypothesis (H_a): A** hypothesis to be considered as an alternate to the null hypothesis.
  - The choice of the alternative hypothesis depends on and should reflect ***the purposes of the hypothesis test.***

$H_a : \mu \neq \mu_0$ - alternate hypothesis test relates to a two-tailed test
$H_a : \mu < \mu_0$ - alternate hypothesis test relates to a left-tailed test (one-tailed)
$H_a : \mu > \mu_0$ - alternate hypothesis test relates to a right-tailed test (one-tailed)

The logic of hypothesis testing
1. Take a random sample.
2. If the data do not provide enough evidence in favor of the alternative hypothesis, do not reject the null hypothesis.
3. If the data provide enough evidence in favor of the alternative, reject the null hypothesis.

9.2 Terms, Errors, and Hypothesis

- **Test statistic:** The statistic used as a basis for deciding whether the null hypothesis should be rejected

- **Rejection region:** The set of values for the test statistic that leads to rejection of the null hypothesis

- **Non-rejection region:** The set of values for the test statistic that leads to non-rejection of the null hypothesis

- **Critical values:** The values of the test statistic that separate the rejection and non-rejection regions

<u>Type I Error</u> & <u>Type II Errors</u>

	H_0 is True	**H_0 is False**
Do not reject H_0	Correct Decision	Type II error
Reject H_0	Type I error	Correct Decision

- **Type I error**: Rejecting the null hypothesis when it is in fact true.

- **Type II error:** Not rejecting the null hypothesis when it is in fact false.

Probabilities of Type I and Type II Errors

- **Significance level α:** The probability of making a Type I error (rejecting a true null hypothesis)
- **Power** of a hypothesis test: Power = 1 – P(Type II error) = 1- β
  - The Probability of rejecting a false null hypothesis
  Power near 0: the hypothesis test is not good at detecting a false null hypothesis.
  Power near 1: the hypothesis is extremely good at detecting a false null hypothesis

Relation between Type I and Type II error probabilities

For a fixed sample size, the smaller we specify the significance level, α, the larger will be the probability, β, of not rejecting a false null hypothesis

Possible conclusions for a hypothesis test

Suppose a hypothesis test is conducted at a small significance level

1. If the null hypothesis is rejected, we conclude that the alternate hypothesis is true.
2. If the null hypothesis is not rejected, we conclude that the data do not provide sufficient evidence to support the alternative hypothesis.

- **Statistically significant at the α level** – null hypothesis is rejected in a hypothesis test performed at the significance level α
- **Not statistically significant at the α level** – null hypothesis is not rejected in a hypothesis test performed at the significance level α

9.3 Hypothesis Tests for one Population Mean when σ is known

Obtaining critical values

Suppose a hypothesis test is to be performed at a specified significance level, α. Then the critical value(s) must be chosen so that, if the null hypothesis is true, the probability is α that the test statistic will fall in the rejection region.

Hypothesis Tests for a Population Mean when σ is known

Procedure 9.1: The one-sample z-test for a population mean

ASSUMPTIONS

1. Normal population or large samples
2. σ is known

STEPS

1. The null hypothesis is $H_0 : \mu = \mu_0$ and the alternate hypothesis is one of the following:

   $H_a : \mu \neq \mu_0$ (two-tailed)
   $H_a : \mu < \mu_0$ (left-tailed)
   $H_a : \mu > \mu_0$ (right-tailed)

2. Decide on the significance level, α.

3. Compute the value of the test statistic $z = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$

4. Choose the critical value(s) according to the alternate hypothesis:
   $\pm z_{\alpha/2}$ (two-tailed)
   $- z_{\alpha}$ (left-tailed)
   z_{α} (right-tailed)

5. If the value of the test statistic falls in the rejection region, reject H_0; otherwise, do not reject H_0.

6. Interpret the results of the hypothesis test.

NOTE: The hypothesis test is exact for normal populations and is approximately correct for large samples from non-normal populations.

When to use the z-test

1. For small samples, say, of size less than 15, the z-test should be used only when the variable under consideration is normally distributed or very close to being so.

2. For moderate size samples, say, between 15 and 30, the z-test can be used unless the data contains outliers or the variable under consideration is far from being normally distributed.

3. For large samples, say, of size 30 or more, the z-test can be used essentially without restriction. However, if outliers are present and their removal is not justified, the effect of the outliers on the hypothesis test should be examined; that is, the hypothesis test should be performed twice, once with the outliers retained and once with them removed. If the conclusion remains the same either way, we may be content to take that as our conclusion and close the investigation. But if the conclusion is affected, it is probably wise to make the more conservation, use a different procedure, or take another sample.

4. If outliers are present but their removal is justified and results in a data set for which the z-test is appropriate, then the procedure can be used.

- **Statistical Significance vs. Practical Significance**
  Statistical significance does not necessary imply practical significance!

- **The Relation Between Hypothesis Tests and Confidence Intervals**

Decision on the level α two-tailed hypothesis test:
Do not reject H_0 if μ_0 is in the $(1 - \alpha)$ confidence interval and reject if outside the interval

9.4 P-Values

Hypothesis Testing:

- **Critical value approach** as in procedure 9.1 (z-test)
- **P-value approach**

Definition 9.5 **P-Value**: To obtain a **P-value** (***P***) of a hypothesis test, we compute, assuming the null hypothesis is true, the probability of observing a value of the test statistic as extreme or more extreme than that observed. By extreme we mean far from what we would expect to observe if the null hypothesis were true.

P-value: referred to as **observed significance level** or probability value

Obtaining P-Values for a One-Sample z-Test

The test statistic for a one-sample z-test for a population mean with null hypothesis $H_0 : \mu = \mu_0$ is

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

If the null hypothesis is true, this test statistics has the standard normal distribution, and its probabilities equal areas under the standard normal curve.
If we let z_0 be the observed value of the test statistic $\boldsymbol{z}$, we obtain the P-value as follows:

- Two-tailed test: The P-value is the probability of observing a value of the test statistic $\boldsymbol{z}$ at least as large in magnitude as the value actually observed, which is the area under the standard normal curve that lies outside the interval from $-|z_0|$ to $|z_0|$.

- Left-tailed test: The P-value is the probability of observing a value of the test statistic $\boldsymbol{z}$ as small as or smaller than the value actually observed, which is the area under the standard normal curve that lies to the left of z_0.

- Right-tailed test: The P-value is the probability of observing a value of the test statistic $\boldsymbol{z}$ as large as or larger than the value actually observed, which is the area under the standard normal curve that lies to the right of z_0.

The P-value approach to Hypothesis Testing

- **P-value as the observed significance level**

The P-value of a hypothesis test is equal to the smallest significance level at which the null hypothesis can be rejected, i.e., the smallest significance level for which the observed sample data results in rejection of H_0.

- **Decision criterion for a hypothesis test using the p-value**

If the P-value is less than or equal to the specified significance level, then reject the null; otherwise, do not reject the null hypothesis.

Procedure 9.2: **The one-sample z-test for a population mean (p-value approach)**

ASSUMPTIONS

1. Normal population or large sample
2. σ is known

STEPS

1. The null hypothesis is $H_0 : \mu \neq \mu_0$ and the alternative hypothesis is one of the following:

   $H_a : \mu \neq \mu_0$ (two-tailed)
   $H_a : \mu < \mu_0$ (left-tailed)
   $H_a : \mu > \mu_0$ (right-tailed)

2. Decide on the significance level, α.

3. Compute the value of the test statistic $z = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$ and denote the value by z_0

4. Use Table II to obtain the P-value

5. If $P \leq \alpha$ reject H_0; otherwise, do not reject H_0.

6. State the conclusion in words.

The hypothesis test is exact for normal populations and is approximately correct for large samples from non-normal populations.

Comparison of the Critical-value and p-value approaches

CRITICAL-VALUE APPROACH	**P-VALUE APPROACH**
1. State H_0 and H_a	1. State H_0 and H_a
2. Decide on α	2. Decide on α
3. Compute test statistic	3. Compute test statistic
4. Determine the critical value(s)	4. determine the P-value
5. If test statistic falls in rejection region, reject H_0; otherwise, do not reject H_0.	5. If $P \leq \alpha$ reject H_0; otherwise, do not reject H_0.
6. Interpret the result of the hypothesis test.	6. Interpret the result of the hypothesis test.

Guidelines for using the P-value to assess the evidence against H_0

P-value	Evidence against H_0
$P > 0.10$	Weak or none
$0.05 < P \leq 0.10$	Moderate
$0.01 < P \leq 0.05$	Strong
$P \leq 0.01$	Very Strong

9.5 Hypothesis Tests for one population mean when σ is unknown

If a variable x of a population is normally distributed with mean μ, then the *studentized* version of $\bar{x}$,

$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$, has t-distribution with n-1 degrees of freedom.

P-value for a t-test

- Two-tailed test: The P-value is the probability of observing a value of the test statistic t at least as large in magnitude as the value actually observed, which is the area under the t-curve that lies outside the interval from $-|t_0|$ to $|t_0|$.

- Left-tailed test: The P-value is the probability of observing a value of the test statistic t as small as or smaller than the value actually observed, which is the area under the t-curve that lies to the left of t_0.

- Right-tailed test: The P-value is the probability of observing a value of the test statistic t as large as or larger than the value actually observed, which is the area under the t-curve that lies to the right of t_0.

Estimating the P-value of a t-test

- Use a computer to get an exact P-value
- Use a Table IV to estimate the range of the P-value

Example: Consider a right-tailed t-test with n=15, α=0.05, and the value of the test statistic of t=3.458.
Solution: For df=15-1=14, the t-value 3.458 is larger than any t-value in Table IV, the largest one being $t_{0.005} = 2.977$. So the P-value is less than 0.005.
The exact P-value from MINITAB is 0.0019.

***Procedure 9.3*: The one-sample t-test (t-test) for a population mean**

ASSUMPTIONS

1. Normal population or large sample
2. σ is unknown

STEPS

1. The null hypothesis is $H_0 : \mu \neq \mu_0$ and the alternative hypothesis is one of the following:

   $H_a : \mu \neq \mu_0$ (two-tailed)
   $H_a : \mu < \mu_0$ (left-tailed)
   $H_a : \mu > \mu_0$ (right-tailed)

2. Decide on the significance level, α.

3. The critical value(s) are

   $\pm t_{\alpha/2}$ (two-tailed)
   $- t_{\alpha}$ (left-tailed)
   t_{α} (right-tailed)

   with df = n – 1. Use Table IV to find the critical value(s).

4. Compute the value of the test statistic

5. If the value of the test statistic falls in the rejection region, reject H_0; otherwise, do not reject H_0.

6. State the conclusion in words.

The hypothesis test is exact for normal populations and is approximately correct for large samples from non-normal populations.

If assumptions are not satisfied, use other techniques including the non-parametric method – the one-sample Wilcoxon signed-rank test.