Unit 3

Hypothesis Testing

Tony Chan, Ph.D.

3.1 Introduction

We conduct ***sample surveys*** in order to ***draw conclusions*** for the ***whole population*** based on ***sample data***.

Example 3.1

There is a common belief that 25% of all the youths in Hong Kong were addictive to drugs in the past.

How to judge if such a claim is believable?

Suppose that 100 youths were randomly sampled from the parent population (All people in Hong Kong), and discovered that 20 of them was addictive to drugs.

Can we conclude that 25% of ALL youths in Hong Kong was drug abusers?

In Unit 2, we have introduced 2 types of parameter estimation methods:

- point estimation
- interval estimation

3.1 Introduction (Cont'd)

Since these 2 methods of estimation only provide ***estimates*** rather than the ***true values*** to the ***population parameters***, therefore error must exist.

What poses the error between a population parameter and an estimate?

- The error may be a sampling error (as in the above example, fewer drug abusers were selected into the sample), or
- the population does not contain 25% or more drug abusers. To address this problem, we can resort to ***hypothesis testing***.

3

3.1 Introduction (Cont'd)

- The main objective of hypothesis testing is to formulate rules that lead to a decision culminating in acceptance or rejection of the statement about the population parameter under test.
- We emphasize on making hypothesis about the population parameter(s). The truth or falsity of a statistical hypothesis is never known with absolute certainty unless we examine the whole population. Instead, we test that hypothesis by looking at a random sample or samples drawn from the population.
- ***Evidence from the sample*** that is ***inconsistent*** with the ***pre-stated hypothesis*** leads to ***rejection of the hypothesis***, whereas ***evidence supporting the hypothesis*** tells us that we have ***insufficient evidence to reject the hypothesis***.

$$\boxed{\text{Population parameter(s)}} \xleftarrow{\text{estimate (infer)}} \boxed{\text{Sample statistic(s)}}$$

4

3.2 Principle of Hypothesis Testing

Example 3.2

Hong Kong International University has a statistical analysis class. According to historical data, student marks follows $N(91, 3^2)$. The marks of a sample of 25 students are displayed below:

87.8, 94.3, 92.2, 95.2, 89.0, 88.2, 87.9, 95.3, 92.1, 96.2, 92.0, 97.0, 93.6, 89.1, 93.9, 91.8, 97.6, 88.4, 90.1, 91.6, 90.7, 93.1, 96.5, 89.8, 90.5

The point estimate of population mean mark is:

$$\hat{\mu} = \bar{x} = \frac{1}{25}(87.8 + \Lambda + 90.5) = 92.2$$

5

3.2 Principle of Hypothesis Testing (Cont'd)

Historical data shows that the marks follow a normal distribution with mean value 91. Although the difference between 92.2 and 91 is small (1.2), *can we still draw the conclusion that they performed better in Mathematics*?

Hypothesis testing is used to justify if the difference between the sample mean mark ($\bar{x}$) and the population mean mark (μ) is due *either* to *sampling error or* to *the actual difference between them.*

If, after hypothesis testing, we can justify that there is a real difference between $\bar{x} = 92.2$ and $\mu = 91$, then we can say that the performance of students in Mathematics was really improved ($\because$ 92.2 > 91).

6

3.2 Principle of Hypothesis Testing (Cont'd)

Example 3.3

According to historical data, the mean number of smokers in Hong Kong in 2013 was 45,000. A certain youth organization claims that the number of smokers is on the rise.

From above, we know that:

Population mean, μ = 45,000. Now, some people raise a new notion:

the number of smokers is increasing.

For the traditional belief, people generally accept μ = 45,000. That is, there is a strong belief for μ = 45,000. However, some other people is doubting the value of μ $(\mu > 45{,}000)$.

3.3 Terminologies

To conduct hypothesis testing, some definitions are needed:

Definition 3.1 - Null Hypothesis

The hypothesis that negates the claim laid under the alternative hypothesis is called the null hypothesis and is denoted by the symbol, H_0. Rejection of this hypothesis may lead to non-rejection of the alternative hypothesis, H_1.

From Example 3.3,

$H_0 : \mu = \mu_0 = 45{,}000$ (the number of smokers remains unchanged)

The null hypothesis is a statement related to the *population parameter*. Here, H_0 is "$\mu = \mu_0 = 45{,}000$" .Such hypothesis is widely accepted, and is not easily rejected. In other words, the event "$\mu = \mu_0 = 45{,}000$" *has a very high probability of occurrence.*

3.3 Terminologies (Cont'd)

Definition 3.2 – **Alternative Hypothesis**

The hypothesis that a researcher believes to be true and expects to establish on the basis of collected data is called the alternative hypothesis or research hypothesis, and is denoted by H_1.

From Example 3.3,

$$H_1 : \mu > \mu_0 = 45{,}000 \text{ (the number of smokers is increasing)}$$

When H_0 is rejected, then H_1 ***may*** be valid.

Exercise 3.1

State the hypotheses, H_0 and H_1, for the following:

1. Is the mean weight of the Hong Kong postmen 1.67m or is it more?

   Let μ = population mean weight of all Hong Kong postmen.

$$\begin{cases} H_0 : \mu = 1.67m \\ H_1 : \mu > 1.67m \end{cases}$$

2. Is the mean weight of the Hong Kong postmen 1.67m or is it something different?

   Let μ = population mean weight of all Hong Kong postmen.

$$\begin{cases} H_0 : \mu = 1.67m \\ H_1 : \mu \neq 1.67m \end{cases}$$

3. On average, do cats have 9 lives, or do they have less than 9?

Let μ = population mean number of lives of all cats.

$$\begin{cases} H_0 : \mu = 9 \\ H_1 : \mu < 9 \end{cases}$$

4. A language center claims that adults will read 400 words/min faster after completing its course of speed reading. A random sample of 24 adults is found to increase their reading speed by 385 words/min. Is the center's claim substantiated?

Let μ = population mean number of words per min after training.

$$\begin{cases} H_0 : \mu = 400 \text{ words/min} \\ H_1 : \mu < 400 \text{ words/min} \end{cases}$$

3.3 Terminologies (Cont'd)

Definition 3.3 – Significance Level (α)

The significance level of a statistical hypothesis test is a fixed probability of wrongly rejecting the null hypothesis H_0 if it is in fact true.

It is the probability of a Type I error and is set by the investigator in relation to the consequences of such an error. That is, we want to make the significance level as small as possible in order to protect the null hypothesis and to prevent, as far as possible, the investigator from inadvertently making false claims.

The significance level is usually denoted by α:

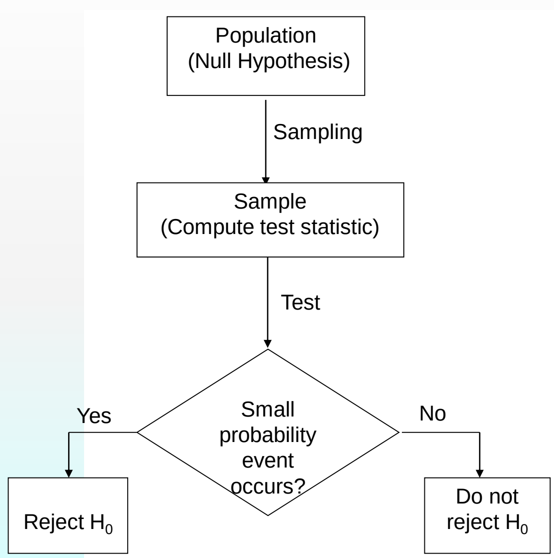
Usually, the significance level is chosen to be 10%, 5%, or 1%.

3.4 Small & Large Probability Events

We give a H_0 with a very high probability (say 95%) of occurrence. This means that the corresponding H_1 has a very low probability (say 5%) of occurrence. If, in a single hypothesis test, the event corresponding to H_1 occurs, then we have ground to doubt the correctness of H_0.

The event corresponds to H_0 is called a *large-probability event*, whereas the event corresponds to H_1 is called a *small-probability event*.

3.5 Flowchart of Hypothesis Testing



3.6 Type I Error & Type II Error

Is our rejection of H_0 always correct?

The rejection of H_0 is based on sample data, and the population parameter under consideration is often unknown. Thus, the inference of population parameter on the basis of sample data may be in doubt. There are two types of error that can be committed:

Definition 3.4 Type I Error - False Positive

Rejection of the H_0 when it is true is called a Type I error. The probability of committing this error is denoted by the Greek letter α, and is referred to as the significance level of the test.

Symbolically,

$$\alpha = P(\text{Type I error})$$
$$= P(\text{Reject } H_0 \mid H_0 \text{ is true})$$

15

Definition 3.5 Type II Error - False Negative

Acceptance of H_0 when it is false is called a Type II error. The probability of making this error is denoted by the Greek letter β.

Symbolically,

$$\beta = P(\text{Type II error})$$
$$= P(\text{Accept } H_0 \mid H_0 \text{ is false})$$

Test Procedure Conclusion	TRUE STATE OF NATURE (UNKNOWN)	
	H_0 is True	H_0 is False
Accept H_0	Correct decision: prob. = 1 - α	Incorrect decision: Type II error; Prob. = β
Reject H_0	Incorrect decision: Type I error; Prob. = α	Correct decision: prob. = 1- β

16

3.7 Statistical Tests for Testing

How to test the correctness of a hypothesis?

In order perform hypothesis testing, we have to devise statistical tests. There are two commonly used tests for hypothesis testing:

Z Test: A test for ***normal population*** mean μ when population SD, σ is known.

t Test: A test for ***normal population*** mean μ when population SD, σ is not known.

Every test of Z test and t test can be divided into three specific tests:

Left-sided (-tailed) test :	$H_0: \mu = \mu_0 \leftrightarrow H_1: \mu < \mu_0$
Right-sided (-tailed) test:	$H_0: \mu = \mu_0 \leftrightarrow H_1: \mu > \mu_0$
Two-sided (-tailed) test:	$H_0: \mu = \mu_0 \leftrightarrow H_1: \mu \neq \mu_0$

3.7.1 Z Test

Z Test

A test for the population mean μ of a normal population when σ^2 is known.

There are 6 steps in performing hypothesis testing:

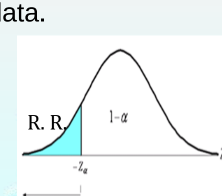
Left-sided Test:

(i) Set up H_0 and H_1 : H_0: $\mu = \mu_0$ $\leftrightarrow$ H_1: $\mu < \mu_0$.

(ii) Given significance level α, choose appropriate Z_α, then determine the rejection region (R.R.) for H_0: $Z < -Z_\alpha$

(iii) Compute sample mean $\bar{x}$ based on the sample data.

(iv) Calculate value of test statistic: $z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$



(v) Make decision: If $z < -z_\alpha$, then reject H_0;
if $z \geq -z_\alpha$, then do not reject H_0 .

(vi) Draw conclusion.

3.7.1 Z Test (Cont'd)

Example 3.4 (Left-tailed Test)

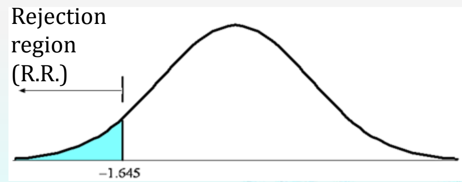
According to past experience, the marks of a certain exam follow $N(70, 6^2)$. The mean of a random sample of 36 students is calculated as 68.5. Test, at 5% level, whether the sample information conforms to the past experience.

(i) H_0: $\mu = 70 \leftrightarrow H_1$: $\mu < 70$

(ii) Take $\alpha = 5\%$, $z_{0.05} = 1.645$. The R.R. is $Z < -1.645$.

(iii) $\bar{x} = 68.5$

(iv) $z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{68.5-70}{6/\sqrt{36}} = -1.5$



(v) $\because z > -1.645$,

$\therefore$ do not reject H_0 at the 5% level.

(vi) At the 5% level, there is no evidence that the sample information does not conform to past experience.

19

3.7.1 Z Test (Cont'd)

Right-sided Test:

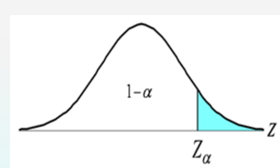
(i) Set up H_0 and H_1 : H_0: $\mu = \mu_0 \leftrightarrow H_1$: $\mu > \mu_0$

(ii) Given significance level α, choose appropriate Z_α, then determine the rejection region for H_0: $Z > Z_\alpha$

(iii) Compute sample mean $\bar{x}$ based on the sample data.

(iv) Compute value of test statistic:

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



(v) Make decision: If $z > z_\alpha$, then reject H_0;

if $z \leq z_\alpha$, then do not reject H_0.

(vi) Draw conclusion.

20

3.7.1 Z Test (Cont'd)

Example 3.5 (Right-tailed Test)

A random sample of 40 of the price ($) of Mathematics books are as follows:

48.5, 38.3, 39.4, 46.0, 45.8, 39.9, 46.9, 47.8, 43.0, 53.7,
39.1, 54.7, 39.8, 42.9, 44.4, 43.0, 43.3, 43.0, 52.7, 64.4,
39.7, 48.2, 44.4, 43.7, 45.8, 42.9, 55.7, 44.9, 33.1, 57.0,
49.5, 46.1, 67.4, 48.5, 61.1, 34.8, 45.8, 64.2, 53.3, 34.7

Assume that the prices follow $N(\mu, 7.6^2)$. Test, at the 1% level, whether there is sufficient evidence that the old mean price of $43.5 has increased.

21

3.7.1 Z Test (Cont'd)

Solution

(i) H_0: $\mu = 43.5 \quad \leftrightarrow \quad H_1$: $\mu > 43.5$

(ii) Take α = 1%, $z_{0.01} = 2.33$. The rejection region is $Z > 2.33$.

(iii) $\bar{x} = 46.9$ (calculated from sample data)

(iv) $z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{46.9-43.5}{7.6/\sqrt{40}} = 2.83$

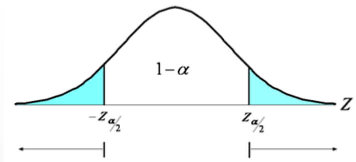
(v) $\because \quad z > 2.33$,

$\therefore$ reject H_0 at the 1% level.

(vi) At the 1% level, there is evidence that the mean price of books has increased.

22

3.7.1 Z Test (Cont'd)



Two-sided Test:

(i) Set up H_0 and H_1 : $H_0: \mu = \mu_0 \quad \leftrightarrow \quad H_1: \mu \neq \mu_0$

(ii) Given significance level α, choose appropriate $Z_{\alpha/2}$, then determine the rejection region for H_0:

(iii) Compute sample mean $\bar{x}$ based on the sample data.

(iv) Compute value of test statistic: $z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$

(v) Make decision: If $Z < -Z_{\alpha/2}$ or $Z > Z_{\alpha/2}$, then reject H_0;

if $-Z_{\alpha/2} \leq Z \leq Z_{\alpha/2}$, then do not reject H_0.

(vi) Draw conclusion.

23

3.7.1 Z Test (Cont'd)

Example 3.6

New City Radio Station doubts that the Rich City residents' listening time for its programs was not equal to 55 minutes. According to past experience, listening time follows $N(\mu, 15^2)$. The researcher of the Station randomly selects a sample of 25 residents in the city, and computes the mean listening time which is 60 minutes. Test, at the 5% level, whether there is evidence that a difference between mean current and past listening times exists.

24

3.7.1 Z Test (Cont'd)

Solution

(i) H_0: $\mu = 55 \;\leftrightarrow\; H_1$: $\mu \neq 55$.
(ii) Take α = 5%, $z_{0.025} = 1.96$.
The rejection region is $Z < -1.96$ or $Z > 1.96$.

(iii) $\bar{x} = 60$
(iv) $z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{60-55}{15/\sqrt{25}} = 1.67$

(v) $\because$ $z < 1.96$,
$\therefore$ do not reject H_0 at the 5% level.

(vi) At the 5% level, there is no evidence that the current mean listening time differs from the past one.

3.7.2 t Test

In practice, σ^2 is usually not known. However, the population data are normally distributed. There are two important points to note:

- Since σ^2 is not known, we need to use sample standard deviation:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

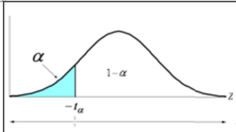
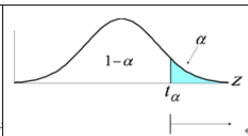
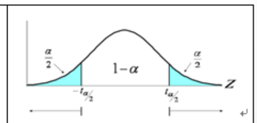
to estimate σ^2.

- *Z* test cannot be used. Rather, we need to use the t test and the test statistic follows the following t distribution:

$$T = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n\text{-}1),$$

where (n-1) is the degrees of freedom.

3.7.2 t Test (Cont'd)

		
Left-sided Test $H_0: \mu = \mu_0 \leftrightarrow H_1: \mu < \mu_0$	Right-sided Test $H_0: \mu = \mu_0 \leftrightarrow H_1: \mu > \mu_0$	Two-sided Test $H_0: \mu = \mu_0 \leftrightarrow H_1: \mu \neq \mu_0$
Rejection Region $T < -t_\alpha$	Rejection Region $T > t_\alpha$	Rejection Region $T < -t_{\alpha/2}$ 或 $T > t_{\alpha/2}$
Example When n = 10, $\alpha = 5\%$, $t_{0.05} = 1.812$	**Example** When n = 10, $\alpha = 1\%$, $t_{0.01} = 2.764$	**Example** When n = 10, $\alpha = 5\%$, $t_{0.025} = 2.22$

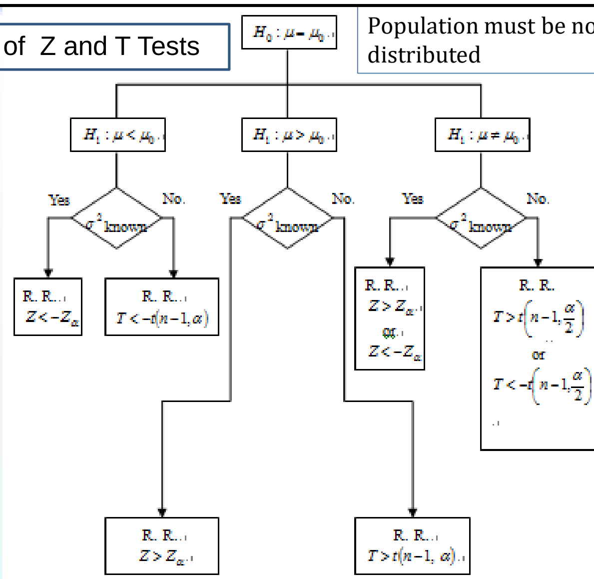
Note:

When the sample size n is at least 30, the t test can be replaced by the Z test.

27

3.7.3 Summary of Z and T Tests

Population must be normal distributed



where

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}, \quad T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

28

Example 3.7

Great Power Motors Factory claims that, under normal conditions, its small motor consumes electricity no more than 0.8 A (unit of electricity), on average. A researcher selects a random sample of 16 motors, and found that, on average, the mean electricity consumption is 0.92 A with standard deviation 0.32 A. Assume that the electricity consumption for this type of motor follows a normal distribution, test, at α = 5%, if the claim of the factory is well-founded.

Example 3.7 (Cont'd)

Since σ^2 is not known and n = 16 < 30, we may resort to T test.

(i) $H_0\colon \mu = \mu_0 = 0.8 \;\leftrightarrow\; H_1\colon \mu > 0.8$

Take = 5%， we have $t_{0.05;15} = 1.753$. The R. R. is $T > 1.753$.

$\bar{x} = 0.92\ (given)$

(ii) $t = \frac{\bar{x}-\mu}{s/\sqrt{n}} = \frac{0.92-0.8}{0.32/\sqrt{16}} = 1.5$

(iii) ∵ t < 1.753,

∴ do not reject H_0 at the 5% level.

(iv) We conclude that, at the 5% level, the claim of the factory is well-founded.

Exercise 3.2

1. John is the manager of a fast-food restaurant. He wants to determine whether or not the waiting time to place an order has changed in the last month from its previous population mean value of 4.5 minutes. From past experience, he can assume that the population standard deviation is 1.2 minutes. He selects a sample of 25 orders during a one-hour period. The sample mean is 5.1 minutes. Determine if there is evidence at the 5% level of significance that the mean waiting time to place an order has changed in the last month from its previous population mean value of 4.5 minutes. Assume that the waiting time distribution is normally distributed.

Solution

(i) $H_0\colon \mu = 4.5 \;\leftrightarrow\; H_1\colon \mu \neq 4.5$

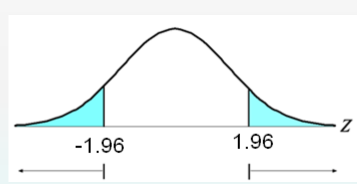
(ii) Take α = 5%, $z_{0.025} = 1.96$

The rejection region is $Z < -1.96$ or $Z > 1.96$

Since population standard deviation is known, therefore Z test is used.

(iii) $\bar{x} = 5.1$

(iv) $z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{5.1-4.5}{1.2/\sqrt{25}} = 2.50$



(v) $\because \; z > 1.96$,

$\therefore$ reject H_0 at the 5% level.

(vi) At the 5% level, there is evidence that the waiting time to place an order has changed from its previous population mean value of 4.5 minutes. In fact, the mean waiting time for customers is longer now than that of last month.

2. The director of manufacturing at a clothing factory needs to determine whether a new machine is producing a particular cloth according to the manufacturer's specifications, which indicates that the cloth should have a mean breaking strength of 70 pounds with a standard deviation of 3.5 pounds. A sample of 49 pieces of cloths reveals a sample mean breaking strength of 69.1 pounds. Determine, at 5% level, if there is evidence that the machine is not meeting the manufacturer's specifications for mean breaking strength.

Solution

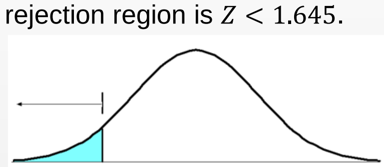
Since population standard deviation is known and the sample size (49) is greater than 30, therefore Z test can be used.

(i) $H_0: \mu = 70 \quad \leftrightarrow \quad H_1: \mu < 70$

(ii) Take α = 5%, $z_{0.05} = 1.645$. The rejection region is $Z < 1.645$.

(iii) $\bar{x} = 69.1$

(iv) $z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{69.1-70}{3.5/\sqrt{49}} = -1.8$



(v) $\because\ z < -1.645$,

$\therefore$ reject H_0 at the 5% level.

(vi) At the 5% level, there is evidence that the machine is not meeting the manufacturer's specifications for mean breaking strength.

3. A manufacturer of flashlight batteries took a sample of 13 batteries from the production of a day and used them continuously until they failed to work. The life times of the batteries in hours until failure are as follows:

342, 426, 317, 545, 264, 451, 1049, 631, 512, 266, 492, 562, 298

At the 5% level of significance, test if there is evidence that the mean life time of the batteries is more than 400 hours. Assume that the life-time distribution is normally distributed.

35

Solution

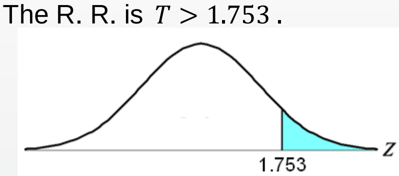
Since the data are normally distributed, σ^2 is not known and n = 13 < 30, we may resort to t test.

(i) H_0: $\mu = 400 \;\leftrightarrow\; H_1$: $\mu > 400$

Take = 5%，we have $t_{0.05;15} = 1.753$. The R. R. is $T > 1.753$.

$\bar{x} = 473.46, s = 210.77$

(iv) $t = \frac{\bar{x}-\mu}{s/\sqrt{n}} = \frac{473.46-400}{210.77/\sqrt{13}} = 1.26$



(v) ∵ t < 1.753,

∴ do not reject H_0 at the 5% level.

(vi) We conclude that, at the 5% level, the mean life time of the batteries is no more than 400 hours.

36

3.7.5 P-value

- P-value answer the question: What is the probability of the observed test statistic … **when H_0 *is true*?**
- Thus, smaller and smaller P-values provide stronger and stronger evidence against H_0
- Small P-value $\Rightarrow$ strong evidence

37

3.7.6 Meaning of "P-value < 0.0001"

Symbolically, "p-value < 0.0001" can be written as

Pr(H_0 is true given the sample data) < 0.0001

- By convention, p-values of < 0.05 are often accepted as "***statistically significant***" in the medical literature. However, this is only an arbitrary cut-off.
- A cut-off of "p < 0.05" means that in about 5 out of 100 experiments, a result would appear significant ***just by chance*** ("Type I error").

38

Example 3.8 (Example 3.7 Revisited: Slide 30)

Critical-value Approach

(i) H_0: $\mu = 0.8 \leftrightarrow H_1$: $\mu > 0.8$
Take = 5%, $t_{0.05;15} = 1.753$.
R. R. is $T > 1.753$.
$\bar{x} = 0.92$ $(given)$

(ii) $t = \frac{\bar{x}-\mu}{s/\sqrt{n}} = \frac{0.92-0.8}{0.32/\sqrt{16}} = 1.5$

(iii) ∵ 1.5 < 1.753,
∴ do not reject H_0 at 5% level.

(iv) We conclude that, at the 5% level, the claim of the factory is well-founded.

P-value Approach

(i) H_0: $\mu = 0.8 \leftrightarrow H_1$: $\mu > 0.8$
Take = 5%.
$\bar{x} = 0.92$ $(given)$

(ii) $t = \frac{\bar{x}-\mu}{s/\sqrt{n}} = \frac{0.92-0.8}{0.32/\sqrt{16}} = 1.5$

(iii) t =1.5, p-value = 0.0772 (obtained by Excel command: "=tdist(t value; d.f.; 1)")
∵ 0.0772 > α = 0.05,
∴ do not reject H_0 at the 5% level.

(iv) We conclude that, at the 5% level, the claim of the factory is well-founded.

3.8 Two-sample Hypothesis Testing

In a two-sample hypothesis test, 2 parameters from 2 populations are compared.

For a two-sample hypothesis test:

- H_0 is a statistical hypothesis that usually states there is ***no difference*** between the parameters of 2 populations. H_0 always contains the symbol "=".
- H_1 is a statistical hypothesis that is true when H_0 is false. H_1 always contains the symbol ">" or "<".

3.8 Two Sample Hypothesis Testing (Cont'd)

To write a null and an alternative hypothesis for a two-sample hypothesis test, translate the claim made about the population parameters from a ***verbal statement*** to a ***mathematical statement***.

Two-tailed Test	Right-tailed Test	Left-tailed Test
$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$	$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 > \mu_2 \end{cases}$	$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 < \mu_2 \end{cases}$

Regardless of which one of the above hypotheses is used, $\mu_1 = \mu_2$ is always assumed to be true.

3.8.1 Two Sample *Z*-Test ($\sigma_1^2 \neq \sigma_2^2$ are known)

Three conditions are necessary to perform a Z-test for the difference between 2 population means μ_1 and μ_2.

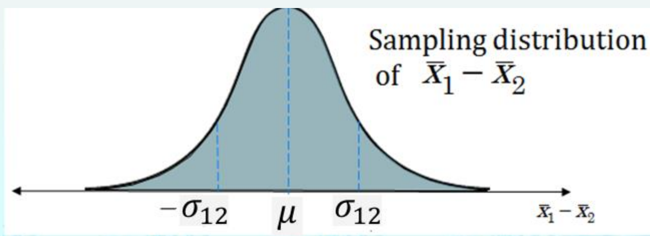
Assumptions:

- The samples must be randomly selected from 2 different pop.
- The samples must be ***independent***. Two samples are ***independent*** if the sample selected from the first population is ***not related*** to the sample selected from the second population.
- Each sample size must be ***at least 30***, or, if not, each population must be ***normally distributed*** with a ***known*** standard deviation.

3.8.1 Two Sample *Z*-Test ($\sigma_1^2 \neq \sigma_2^2$ are known)

If the above assumptions are met, the ***sampling distribution*** of $\bar{X}_1 - \bar{X}_2$ (the difference of the sample means) is a ***normal distribution*** with mean (μ) and standard error (σ) given by the following expressions:

$$\mu = \mu_1 - \mu_2 \quad \text{and} \quad \sigma_{12} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$



43

3.8.1 Two Sample *Z*-Test ($\sigma_1^2 \neq \sigma_2^2$ are known) (Cont'd)

Two-Sample *Z*-Test A two-sample *z*-test can be used to test the difference between 2 population means μ_1 and μ_2 with the following ***assumptions***:

- The samples are independent.
- $\bar{X}_1 - \bar{X}_2$ is normally distributed if $n_1 < 30$ ***and*** $n_2 < 30$, ***or***
- ***Sufficiently large samples (n ≥ 30)*** are randomly selected from ***both populations***

The test statistic is

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{12}}, \text{ where } \quad \sigma_{12} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

44

3.8.1 Two Sample *Z*-Test ($\sigma_1^2 \neq \sigma_2^2$ are known) (Cont'd)

Using a Two-Sample *Z*-Test for the Difference Between Means ***(Large Independent Samples)***

1. State the claim mathematically. Identify H_0 and H_1.
2. Specify the level of significance, α.
3. Sketch the sampling distribution.
4. Determine the critical value(s):
   - one-sided test: $-Z_\alpha$ or Z_α;
   - two-sided test: $-Z_{\alpha/2}\ or\ \ Z_{\alpha/2}$.
5. Determine the critical/rejection region(s).

45

3.8.1 Two Sample *Z*-Test ($\sigma_1^2 \neq \sigma_2^2$ are known) (Cont'd)

Using a Two-Sample *z*-Test for the Difference Between Means (Large Independent Samples)

6. Calculate the value of the test statistic.

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{12}},$$

7. Decide to reject or not to reject H_0.
   - If z is in the rejection region, reject H_0;
   - otherwise, do not reject H_0.
8. Draw conclusion based on the context of the original claim.

46

Example 3.9

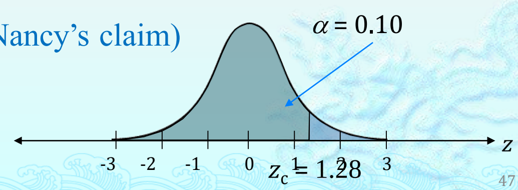
A math teacher, Nancy claims that the students in her class will score higher on the common math exam than the students in another teacher's class. The mean math score for 49 students in her class is 22.1 and the standard deviation is 4.8. The mean math score for 44 of another teacher's students is 19.8 and the standard deviation is 5.4. Test Nancy's claim at the 10% level of significance.

Solution

$\because n_1 = 49 > 30, n_2 = 48 > 30,$

$\therefore$ the sampling distribution of $\bar{X}_1 - \bar{X}_2$ is approximately normally distributed. Z test can be used.

H_0: $\mu_1 = \mu_2$ $\leftrightarrow$ H_1: $\mu_1 > \mu_2$ (Nancy's claim)



$\because n_1 = 49 > 30, n_2 = 44 > 30,$

$\therefore$ Z test can be used.

The standard error is

$$\sigma_{12} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{4.8^2}{49} + \frac{5.4^2}{44}} \approx 1.0644$$

The test statistic, under H_0, is

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{12}} = \frac{(22.1 - 19.8) - (0)}{1.0644} = 2.161$$

$\because 2.161 > z_{0.1} = 1.28,$

$\therefore$ Reject H_0 at the 10% level.

There is enough evidence at the 10% level to support Nancy's claim that her students performed better in the exam.

3.8.2 Two Sample *Z*-Test ($\sigma_1^2 = \sigma_2^2 = \sigma^2$ are known)

Two-Sample *Z*-Test

A two-sample *Z*-test can be used to test the difference between 2 population means μ_1 and μ_2 with the following ***assumptions***:

- The samples are ***independent***.
- $\bar{X}_1 - \bar{X}_2$ is ***normally distributed*** if $n_1 < 30$ and $n_2 < 30$, ***or***
- ***Sufficiently large samples (n ≥ 30)*** are randomly selected from ***both populations***

The test statistic is

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{12}}, \text{ where}$$

$$\sigma_{12} = \sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}} = \sqrt{\sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sigma\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \quad \text{and } \sigma \text{ is given.}$$

49

Exercise 3.3

1. For a sample of **45** adult European males picked at random, the mean weight was 154 pounds, whereas for a sample of **49** adult males in the United States, the mean weight was 162 pounds. Assume that the population variance of weight in Europe is 100 and that in the United States is 169. Is it true that there is a significant difference between mean weights in the two places? Use α = 0.05. Assume that the weights are normally distributed.

50

Solution

$\because$ Both sample sizes $n_1 = 45 > 30$ and $n_2 = 49 > 30$,
$\therefore$ Z test can be used.
H_0: $\mu_1 = \mu_2$ or $\mu_1 - \mu_2 = 0$ $\leftrightarrow$ H_1: $\mu_1 \neq \mu_2$

$\alpha = 0.05 \Rightarrow z_{0.05} = 1.96$

Decision rule: Reject if $Z < -1.96$ or $Z > 1.96$

$\overline{x_1} = 154$, $\overline{x_2} = 162$, $100 = \sigma_1^2 \neq \sigma_2^2 = 169$ (pop variances are <u>not</u> equal)

$$z = \frac{\overline{x_1} - \overline{x_2} - 0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{154 - 162}{\sqrt{\frac{100}{45} + \frac{169}{49}}} = -3.36$$

$\because$ -3.36 < -1.96,
$\therefore$ reject H_0 at the 5% level.

There is a difference between the true mean weight of male adults in the United States and those in Europe.

2. For a sample of **45** adult European males picked at random, the mean weight was 154 pounds, whereas for a sample of **49** adult males in the United States, the mean weight was 162 pounds. Assume that the population variances of weight in Europe and the United States are both equal to 145 pounds. Is it true that there is a significant difference between mean weights in the two places? Use α = 0.05. Assume that the weights are normally distributed.

Solution

$\because$ Both sample sizes $n_1 = 45 > 30$ and $n_2 = 49 > 30$,
$\therefore$ Z test can be used.
H_0: $\mu_1 = \mu_2$ or $\mu_1 - \mu_2 = 0$ $\leftrightarrow$ H_1: $\mu_1 \neq \mu_2$

$\alpha = 0.05 \Rightarrow z_{0.05} = 1.96$

Decision rule: Reject if $Z < -1.96$ or $Z > 1.96$

$\overline{x_1} = 154, \overline{x_2} = 162,$ $\sigma_1^2 = \sigma_2^2 = 145$ (pop variances are equal)

$$z = \frac{\overline{x_1} - \overline{x_2} - 0}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{154-162}{\sqrt{145\left(\frac{1}{45} + \frac{1}{49}\right)}} = -3.2177$$

$\because$ -3.2177 < -1.96,
$\therefore$ reject H_0 at the 5% level.

There is a difference between the true mean weight of male adults in the United States and those in Europe.

Testing Difference Between Means

(Small Independent Samples):

Unpaired t-test

3.9 Unpaired t-test

If 2 samples of sizes ***< 30*** are taken from 2 ***normally distributed populations of unknown variances***, a t-test may be used to test the difference between the population means μ_1 and μ_2.

Assumptions for use a t-test for ***small independent*** samples:

1. The samples must be ***randomly selected from 2 different pop***.
2. The samples must be ***independent***.
3. Each population must be ***normally distributed***.

55

3.9 Unpaired t-test (Cont'd)

Unpaired t-test for the Difference between Means

A two-sample t-test is used to test the difference between two population means μ_1 and μ_2. The test statistic is

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sigma_{12}}$$

Case A: Population variances are EQUAL but UNKNOWN

(i) Combine/Pool sample sizes (n_1, n_2) and sample variances (s_1^2, s_2^2) to estimate σ_{12}.

(ii) Denote the estimate as $\hat{\boldsymbol{\sigma}}_{\mathbf{12}}$.

$$\hat{\sigma}_{12} = S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, \qquad \text{where} \qquad S_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

and d.f.$= n_1 + n_2 - 2$

56

3.9 Unpaired t-test (Cont'd)

The standard error ($\hat{\boldsymbol{\sigma}}_{12}$) for the sampling distribution of $\overline{X_1} - \overline{X_2}$:

Case B: **Unequal and unknown** population variances: $\sigma_1^2 \neq \sigma_2^2$

(i) Directly use s_1^2 and s_2^2 to replace $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$, respectively

$$\hat{\boldsymbol{\sigma}}_{12} = \sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

(ii) d.f. = $min(n_1 - 1, n_2 - 1)$

57

3.9.1 Normal Distribution or t-distribution?

Are both sample sizes ≥ 30? → **Yes** → Use the *Z*-test.

↓ **No**

Are both populations normally distributed? → **No** → You cannot use the *Z*-test or the *t*-test.

↓ **Yes**

Are both population standard deviations known? → **No** → Are the population variances equal?

Are the population variances equal? → **Yes** → Use the t-test with $\hat{\sigma}_{12} = S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$, d.f. = $n_1 + n_2 - 2$

Are the population variances equal? → **No** → Use the t-test with $\hat{\sigma}_{12} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$, d.f. = $min(n_1 - 1, n_2 - 1)$

Are both population standard deviations known? → **Yes** → Use the *Z*-test.

$where\ s.e.(\overline{X_1} - \overline{X_2}) = \sigma_{12}$

58

3.9.2 Procedures of Unpaired t-test

Using a Unpaired t-test for the Difference between Means (*Small Independent Samples*)

1. State the claim mathematically. Identify H_0 and H_1.
2. Specify the level of significance, α.
3. Identify the d.f. ($n_1 + n_2 - 2$ or $\min(n_1 - 1, n_2 - 1)$)
4. Sketch the critical region of the sampling distribution.
5. Determine the critical value(s).

3.9.2 Procedures of Unpaired t-test (Cont'd)

Using a Unpaired t-test for the Difference between Means (*Small Independent Samples*)

5. Determine the rejection region(s).
6. Compute the test statistic:

$$t = \frac{(\overline{X_1} - \overline{X_2}) - (\mu_1 - \mu_2)}{\hat{\sigma}_{12}}$$

7. Decide to reject or fail to reject H_0.
   - If t is in the rejection region, reject H_0;
   - otherwise, do not reject H_0.
8. Draw conclusion based on the context of the original claim.

Example 3.10

A random sample of 17 nurses in Hong Kong has a mean monthly income of \$35,800 and a standard deviation of \$7,800. In Macau, a random sample of 18 nurses has a mean monthly income of \$35,100 and a standard deviation of \$7,375. Test the claim at $\alpha = 0.01$ that the mean monthly incomes in the two cities are ***not*** the same. Assume the population variances are equal but unknown and the monthly incomes are normally distributed.

Solution

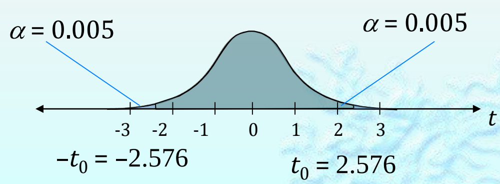
$\because n_1 = 17 < 30, n_2 = 18 < 30$ and the pop variances are equal but unknown,

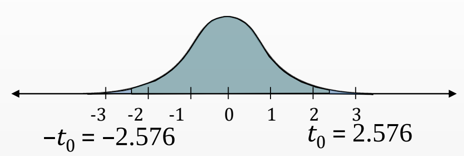
$\therefore$ t test is used.

$H_0: \mu_1 = \mu_2 \quad \leftrightarrow \quad H_1: \mu_1 \neq \mu_2$

d.f. $= n_1 + n_2 - 2$

$= 17 + 18 - 2 = 33$

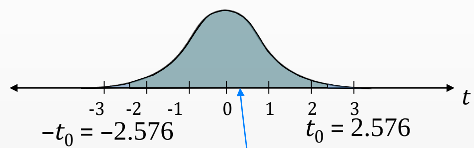




The standard error is

$$\hat{\sigma}_{12} = S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$= \sqrt{\frac{(17 - 1)7800^2 + (18 - 1)7375^2}{17 + 18 - 2}}\sqrt{\frac{1}{17} + \frac{1}{18}} \approx 2564.92$$



Under H_0, the test statistic is

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\hat{\sigma}_{12}} = \frac{(35800 - 35100) - 0}{2564.92} \approx 0.273$$

$\because$ 0.273 < 2.576,

$\therefore$ do not reject H_0 at the 1% level.

At the 1% level, there is insufficient evidence to support the claim that the mean monthly incomes differ.

Example 3.11

The president of a research company is interested in examining the annual mean salary differences between vice-presidents of banks and vice-presidents of loans firms. A random sample of 8 of each kind of vice-president was selected. Their annual salaries (US$) are as follows:

n	1	2	3	4	5	6	7	8
Bank	84320	67340	98590	111780	48940	56790	77610	62000
Loans Firm	73420	49580	58750	101400	88670	59640	65590	74810

Conduct a test of hypothesis to determine if there is a significant difference in the average salary for the 2 vice-president groups. The salaries for both groups are considered to be approximately normally distributed. Use a significance level of 5%. Do not assume that the population variances are equal.

Solution

$\bar{x}_b = \$75921.25$, $s_b = \$21496.00$

$\bar{x}_l = \$71482.50$, $s_l = \$16996.20$

Since the sample sizes are small (< 30) and the samples are collected from 2 different approximately normal populations, unpaired t tests can be used.

H_0: $\mu_1 = \mu_2$ $\quad\leftrightarrow\quad$ H_1: $\mu_1 \neq \mu_2$,

where μ_1 and μ_2 are the mean salaries of vice-presidents of banks and loans firms, respectively.

$$\hat{\sigma}_{12} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{21496.00^2}{8} + \frac{16996.20^2}{8}} = 9688.58$$ (Assume unequal variances)

d.f. = $min(8-1, 8-1) = 7$

Under H_0, the test statistic is

$$t = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\hat{\sigma}_{12}} = \frac{(75921.25 - 71482.50) - (0)}{9688.58} = 0.45 < t(0.025; 7) = 2.365$$

We do not reject H_0 at 5% level and conclude that there is no significant difference between the mean salaries of vice-presidents of the banks and loans firms.

Exercise 3.4

1. A nitrogen fertilizer was used on 10 plots and the mean yield per plot $\overline{x_1}$ was found to be 82.5 bushels with s_1 = 10 bushels. On the other hand, 15 plots treated with phosphate fertilizer gave a mean yield $\overline{x_2}$ of 90.5 bushels per plot with $s_2 = 20$ bushels. At the 5% level of significance, are the 2 fertilizers significantly different? It may be assumed that the yields for the two fertilizers are normally distributed with unequal variances.

Solution

$H_0 : \mu_1 = \mu_2 \qquad H_1 : \mu_1 \neq \mu_2.$

$\because n_1 = 10 < 30, n_2 = 15 < 30$, and σ is unknown and unequal,

$\therefore$ the T test is used.

$n_1 = 10, n_2 = 15, n_1 + n_2 - 2 = 23, \alpha = 5\% \Rightarrow t(23;0.025) = 2.069.$

Decision rule: Reject if $T_0 < -2.069$ or $T_0 > 2.069.$

$$\sigma_{12} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{10^2}{10} + \frac{20^2}{15}} = 6.0553$$

Under H_0, the test statistic is

- $t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{12}} = \frac{(82.5 - 90.5)}{6.0553} = -1.4863$

$\because$ -1.4863 > -2.069,

$\therefore$ do not reject H_0 at the 5% level.

The data do not support the contention that the fertilizer are significantly different.

67

(2) The production supervisor of a factory is conducting a test of the tensile strengths of 2 types of copper coils. The relevant data are as follows:

Coil	Sample Mean	Sample Standard Deviation	Sample Size
A	118	17	9
B	143	24	16

The tensile strengths for the 2 types of copper coils are approximately normally distributed. Do the sample data support the conclusion that the mean tensile strengths of the 2 coils are different at a significance level of 5%? Assume that the population variances of tensile strengths are equal.

68

Solution

$\because n_1 = 9 < 30, n_2 = 16 < 30$ and the pop variances are equal but unknown,

$\therefore$ t test is used.

H_0: $\mu_1 = \mu_2$ $\leftrightarrow$ H_1: $\mu_1 \neq \mu_2$

d.f. $= n_1 + n_2 - 2 = 9 + 16 - 2 = 23$

$$S_p = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}} = \sqrt{\frac{8\times17^2+15\times24^2}{23}} = 21.8214 \text{ (Assume equal variances)}$$

$$\hat{\sigma}_{12} = S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} = 21.8214\sqrt{\frac{1}{9}+\frac{1}{16}} = 9.0923 \text{ (Assume equal variances)}$$

Under H_0, the test statistic is

$$t = \frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\hat{\sigma}_{12}} = \frac{(118-143)-(0)}{9.0923} = -2.7496 < t(0.025;23)\text{=-2.069},$$

We reject H_0 at 5% level and conclude that there is significant difference between the mean tensile strengths of the 2 types of coils.

69

Testing Difference between Means:

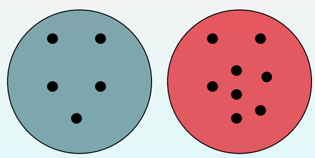
Paired t-test

(*Dependent Samples*)

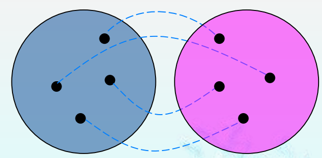
70

3.10 Independent & Dependent Samples

Two samples are ***independent*** if the sample selected from one population is not related to the sample selected from the second population. Two samples are ***dependent*** if each member of one sample corresponds to the ***same member*** of the other sample. Dependent samples are also called ***paired samples*** or ***matched samples.***



Independent Samples



Dependent Samples

3.10 Independent & Dependent Samples (Cont'd)

Example 3.11
Classify each pair of samples as independent or dependent.

Sample 1: The weight of 24 students in a first-grade class
Sample 2: The height of the same 24 students

These samples are dependent because the weight and height can be paired with respect to each student.

Sample 1: The average price of 15 new trucks
Sample 2: The average price of 20 used vans

These samples are independent because it is not possible to pair the new trucks with the used vans. The data represents prices for different vehicles.

3.11 Paired t-test

The following symbols are used for the paired t-test for μ_d:

Symbol	Description
n	Number of pairs of data
d	Difference between entries for a data pair, $d_i = x_{1i} - x_{2i}$
μ_d	Population mean of the differences of paired data
$\bar{d}$	Sample mean of the differences between the paired data $$\bar{d} = \frac{\sum d}{n}$$
s_d	The SD of the differences between the paired data: $$s_d = \sqrt{\frac{n(\sum d^2) - (\sum d)^2}{n(n-1)}}$$

3.11 Paired t-test (Cont'd)

To perform a two-sample hypothesis test with ***dependent*** samples, the difference between each data pair is firstly calculated:

$$d_i = x_{1i} - x_{2i}, \qquad i = 1,2,\cdots,n.$$

The test statistic is the mean $\bar{d}$ of these differences.

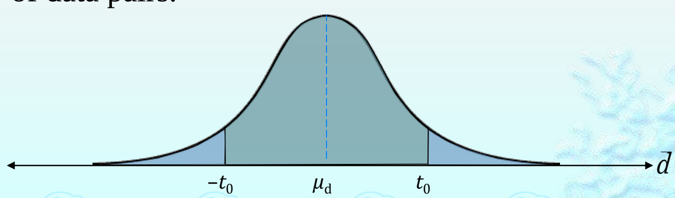
$$\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i$$

Three assumptions must be satisfied to conduct the test.

3.11 Paired t-test (Cont'd)

1. The samples must be randomly selected.
2. The samples must be dependent (paired).
3. Both populations must be normally distributed.

If these assumptions hold, then the sampling distribution for $\bar{d}$ can be approximated by a *t*-distribution with (*n* – 1) d.f., where *n* is the the number of data pairs.



75

3.11 Paired t-test (Cont'd)

Paired t-test for the Difference between Means (*Dependent Samples*)

A t-test can be used to test the difference of 2 population means when a sample is randomly selected from each population. The test statistic is

$$t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}}.$$

The degrees of freedom are (*n* – 1).

76

3.12 Procedures of Conducting a Paired t-test

Using Paired t-test for the Difference between Means (*Dependent Samples*)

1. State the claim mathematically. Identify H_0 and H_1.
2. Specify the level of significance, α.
3. Identify the d.f. (n-1)
4. Sketch the critical region of the sampling distribution.
5. Determine the critical value(s).

77

3.12 Procedures of Conducting a Paired t-test (Cont'd)

Using a Paired t-test for the Difference between Means (*Dependent Samples*)

5. Determine the rejection region(s).
6. Calculate $\bar{d}$ and s_d:

$$\bar{d} = \frac{\sum d}{n}$$

$$s_d = \sqrt{\frac{n(\sum d^2) - (\sum d)^2}{n(n-1)}}$$

7. Calculate the test statistic.

$$t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}}$$

78

3.12 Procedures of Conducting a Paired t-test (Cont'd)

Using a Paired t-test for the Difference between Means (*Dependent Samples*)

8. Decide to reject or not to reject H_0
   - If t is in the rejection region, reject H_0;
   - otherwise, do not reject H_0.

9. Draw conclusion based on the context of the original claim.

Example 3.12

A tuition center claims that students will perform better on an English test after going through the English course offered by the center. The table below shows the scores of 6 students ***before and after*** the course. At $\alpha = 0.05$, is there enough evidence to conclude that the students' scores after the course are better than those before the course? Assume that the scores are normally distributed.

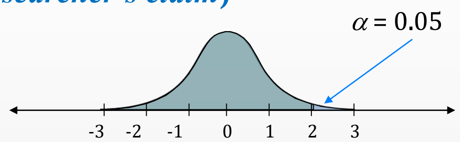
Student	1	2	3	4	5	6
Score (before)	85	96	70	76	81	78
Score (after)	88	85	89	86	92	89

$\because n_1 = 6 < 30, n_2 = 6 < 30,$ *and* σ *is unknown,*

$\therefore$ the T test is used.

H_0: $\mu_d = 0 \quad \leftrightarrow \quad H_1$: $\mu_d > 0$ **(*Researcher's claim*)**

d.f. = 6 – 1 = 5



d = (score before) – (score after)

$t_0 = 2.015$

Student	1	2	3	4	5	6
Score (before)	85	96	70	76	81	78
Score (after)	88	85	89	86	92	89
d	–3	11	–19	–10	–11	–11
d^2	9	121	361	100	121	121

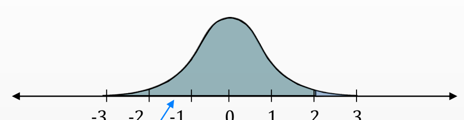
$\sum d = -43$

$\sum d^2 = 833$

$$\bar{d} = \frac{\sum d}{n} = \frac{-43}{6} \approx -7.167$$

$$s_d = \sqrt{\frac{n(\sum d^2) - (\sum d)^2}{n(n-1)}} = \sqrt{\frac{6(833) - 1849}{6(5)}} \approx \sqrt{104.967} \approx 10.245$$

81



$t_0 = 2.015$

The test statistic is

$$t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \frac{-7.167 - 0}{10.245/\sqrt{6}} \approx -1.714.$$

$\because$ -1.714 < 2.015,

$\therefore$ do not reject H_0 at the 5% level.

There is not enough evidence at the 5% level to support the claim that the students' scores after the course are better than those before the course.

82

Exercise 3.5

1. An experiment was performed to compare the abrasive wear of 2 different laminated materials. Twelve pieces of material I were tested by exposing each piece to a machine measuring wear. Ten pieces of material II were similarly tested. In each case, the depth of wear was observed. The samples of material I gave an average wear of 85 units with a sample standard deviation of 4, whereas the samples of material II gave an average of 81 and a sample standard deviation of 5. Test, at the 5% level of significance, if the abrasive wear of material II exceeds that of material II by more than 2 units.

Assume the populations are approximately normally distributed with equal variances.

Solution

Let μ_i = population mean of the abrasive wear for material i, $i = 1, 2$.

$\because n_1 = 12 < 30, n_2 = 10 < 30, and$ σ is unknown, ∴ t test is used.

$$H_0 : \mu_1 - \mu_2 = 2 \quad \leftrightarrow \quad H_1 : \mu_1 - \mu_2 > 2.$$

$$\bar{x}_1 = 85, s_1 = 4, n_1 = 12,$$

$$\bar{x}_2 = 81, s_2 = 5, n_2 = 10.$$

$\alpha = 5\%, t_{0.05;20} = 1.725$

Critical region: $T > t_{\alpha;n_1+n_2-2} = 1.725.$

Since the population variances are unknown but equal, therefore

$$S_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} = \sqrt{\frac{11\times 16 + 9\times 25}{12+10-2}} = 4.4777$$

$$\hat{\sigma}_{12} = S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} = 4.4777\sqrt{\frac{1}{12}+\frac{1}{10}} = 1.9172$$

$$t = \frac{(85-81)-2}{1.9172} = 1.0432 < 1.725$$

Do not reject H_0 at the 5% level. The abrasive wear of material I does not exceed that of material II by more than 2 units.

(2) The market research staff at Allied Foods is considering 2 different packaging designs for an instant breakfast cereal that Allied is about to introduce. The first type of container under consideration is a rectangular box, whereas the second container type has a cylindrical shape. The staff decides to conduct a pilot study by placing the product in both containers and locating the 2 types at opposite ends of the breakfast cereal section in 10 different supermarkets. All the container are placed at eye level to remove any effect due to the height of the display. The main question under consideration is whether there is any difference in the sales of the 2 types of containers. From the data, can we conclude that there is a difference in sales for the rectangular and cylindrical containers? Use α = 5% to define the test.

Supermarket	1	2	3	4	5	6	7	8	9	10
Rectangular	194	152	160	172	118	110	137	126	176	145
Cylindrical	184	161	153	184	105	123	155	111	156	129

Solution

Let μ_d = the average difference in sales for the 2 container types.

Supermarket	1	2	3	4	5	6	7	8	9	10	Total
d	10	-9	7	-12	13	-13	-18	15	20	16	29
d^2	100	81	49	144	169	169	324	225	400	256	1917

$$H_0:\ \mu_d = 0 \quad \leftrightarrow \quad H_1:\ \mu_d \neq 0$$

Reject H_0 if $|t_D| > t_{\alpha/2;n-1} = t_{0.025;9} = 2.262$.

The sample mean and standard deviation of the above data are: $\bar{d} = 2.9, s_d = 14.271$

Under H_0, the test statistic is: $t_d = \frac{\bar{d}}{s_d/\sqrt{n}} = \frac{2.9}{14.271/\sqrt{10}} = 0.643$

∵ 0.643 < 2.262,

∴ we do not reject H_0 at 5% level.

We conclude that there is insufficient evidence to conclude that the container type has an effect on sales.