CHAPTER 7

USING SAMPLE STATISTICS TO TEST HYPOTHESES ABOUT POPULATION PARAMETERS

- **Key words:**

- ✓ Null hypothesis H_0
- ✓ Alternative hypothesis H_A
- ✓ Testing hypothesis
- ✓ Type I error α, Type II error β, Power of the test $(1-\beta)$
- ✓ Test statistic
- ✓ Testing hypotheses concerning one population mean
- ✓ Testing hypotheses concerning two population means
- ✓ Paired comparison
- ✓ P-value

7.1 INTRODUCTION

- **Definition of a hypothesis**
- It is a statement or claim or assertion about one or more populations. It is usually concerned with the parameters of the population. e.g.,

✓ The hospital administrator may want to test the hypothesis that the average length of stay of patients admitted to the hospital is 5 days.

✓ A physician may hypothesize that a certain drug will be effective in 90% of cases for which it is used.

- **Types of Hypotheses**
- They are two types of hypotheses:
- Research hypothesis is the conjecture or supposition that motivates the research.
  - For example, a physician may recall numerous instants in which certain combination of therapeutic measures were more effective than any one of them alone.
- Statistical hypotheses are hypotheses that are stated in such a way that they may be evaluated by appropriate statistical techniques.

There are two statistical hypotheses involved in hypothesis testing

- **Null hypothesis** H_0: It is the hypothesis to be tested.
  - To conduct a hypothesis test about the mean of a population, we postulate a value for it, and call that value μ_0.
  - We write this $H_0: \mu = \mu_0$
- **Alternative hypothesis** H_A: It is a statement of what we believe is true if our sample data cause us to reject the null hypothesis.
  - We set an alternative hypothesis for all other values of μ
  - We write this: $H_A: \mu \neq \mu_0$

Result of the test	Condition of Null Hypothesis	
	$\mu = \mu_0$ (H_0 is true)	$\mu \neq \mu_0$ (H_0 is False)
Do Not Reject H_0	**Correct Decision**	**Type II error (β)**
Reject H_0	**Type I error (α)**	**Correct Decision**

α = Probability of Type I Error
= P(rejecting H_0 given H_0 is true)

β = Probability of Type II Error
= P(Accepting H_0 given H_A is true)

(1- β) = Power of a test
= P(Rejecting H_0 given H_A is true)

Simple and composite hypotheses

H_0: $\mu = 70$ this is simple hypothesis

H_0: $\mu \leq 70$ this is composite hypothesis

H_A: $\mu \geq 70$ this is composite hypothesis

One sided and two sided hypotheses

H_0: $\mu=70$ vs H_A: $\mu > 70$ one sided hypothesis (or one tailed)

H_0: $\mu=70$ vs H_A: $\mu < 70$ one sided hypothesis (or one tailed)

H_0: $\mu=70$ vs H_A: $\mu \neq 70$ two sided hypothesis (or two tailed)

H_0: $\mu=70$ vs H_A: $\mu > 70 \Leftrightarrow H_0$: $\mu \leq 70$ vs H_A: $\mu > 70$

H_0: $\mu=50$ vs H_A: $\mu < 50 \Leftrightarrow H_0$: $\mu \geq 50$ vs H_A: $\mu < 50$

Suppose we want to test

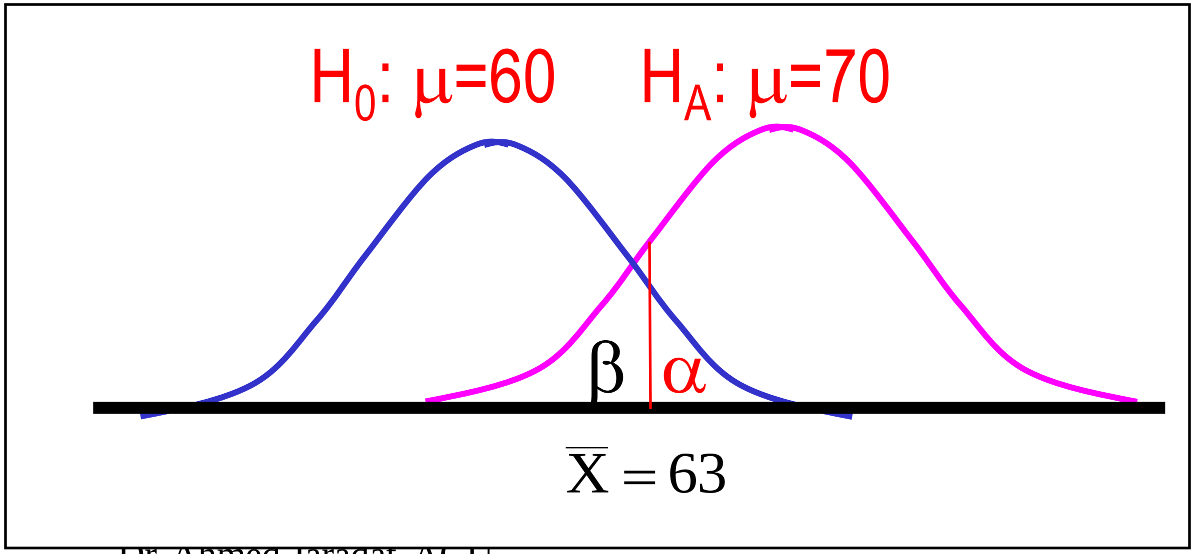
H_0: μ=60 against H_A: μ=70

If we decide to reject H_0 if $\overline{X} > 63$, then

$$\alpha = P(\overline{X} > 63 | \mu = 60)$$

$$\beta = P(\overline{X} < 63 | \mu = 70)$$

$$\text{Power} = 1 - \beta = P(\overline{X} > 63 | \mu = 70)$$



Type I and Type II errors

- Type I error: α is the specified significance level
- Type II: β generally **unspecified** and **unknown**
- **For a given n, α and β inversely related**
- Both types of errors may be reduced **simultaneously** by increasing n
- P-value is the smallest value of α for which we can reject a null hypothesis.
  - ***Reject H_0 if P-value less than α.***

Steps of Hypothesis Testing

1. State the statistical hypotheses, H_0 and H_A.
2. Select a level of significance, α.
3. Find the critical value(s).
4. Compute the test statistics.
5. Refer to a criterion for evaluating the sample evidence.
6. Make a decision to keep/reject the null.
7. Make your comments and conclusion.

7.2 HYPOTHESES TESTING: A SINGLE POPULATION MEAN

A: Sampling From Normally Distributed Population with Known σ (n small or large)

Null hypothesis H_0: $\mu = \mu_0$, α=level of significance

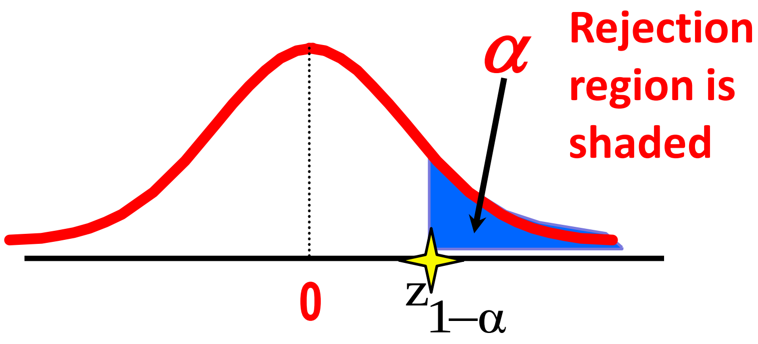
Test statistics: $Z_{cal} = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$

a)

$$\text{Alternative } H_A : \mu > \mu_o$$

$$\text{Reject } H_0 \text{ if } Z_{cal} > z_{1-\alpha}$$

$$P - \text{value} = P\left(Z > Z_{cal}\right)$$

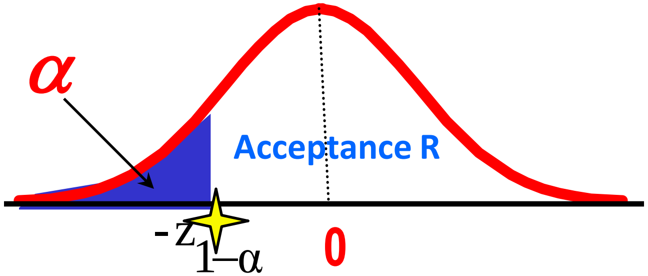


b)

$$\text{Alternative } H_A: \mu < \mu_o$$

$$\text{Reject } H_0 \text{ if } Z_{cal} < -z_{1-\alpha}$$

$$P-\text{value} = P\left(Z < Z_{cal}\right)$$



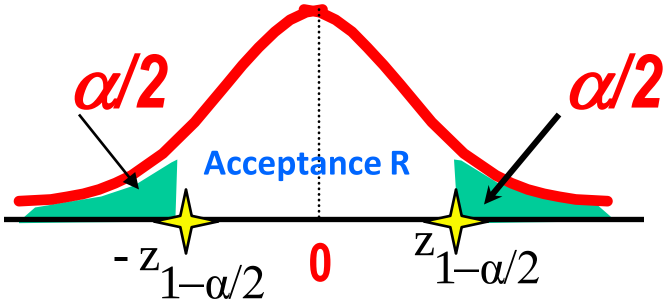
Represents critical value

c)

$$\text{Alternative } H_A: \mu \neq \mu_o$$

$$\text{Reject } H_0 \text{ if } \left|Z_{cal}\right| > z_{1-\alpha/2}$$

$$P-\text{value} = 2P\left(Z > \left|Z_{cal}\right|\right)$$



Example: 7.2.1 page 223

- Researchers are interested in the mean age of a certain population.
- A random sample of 10 individuals drawn from the population of interest has a mean of 27.
- Assuming that the population is approximately normally distributed with variance 20.
- Can we conclude that the mean is **different from 30** years? ($\alpha=0.05$)
- Find the p–value.

Solution-Example: 7.2.1 page 223

- H_0: $\mu=30$ vs H_A: $\mu \neq 30$
- Critical value=$Z_{1-\alpha/2}=Z_{0.975}=1.96$
- Test statistics: $z_{cal}=\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}=\dfrac{27-30}{\sqrt{20}/\sqrt{10}}=-2.12$
- Reject H_0 *since* $z = |-2.12| > 1.96$
- P-value $=2P(Z > |-2.12|)=2(0.0174)=0.0348$
- We conclude that μ is not equal to 30.
- **95%confidence interval for μ is given by:**
- $\left[\bar{X}-z_{1-\alpha/2}\dfrac{\sigma}{\sqrt{n}},\ \ \bar{X}+z_{1-\alpha/2}\dfrac{\sigma}{\sqrt{n}}\right]$ **=[24.229, 29.771]**

Example: 7.2.2 page 229

Referring to example 7.2.1. Suppose that the researchers have asked: Can we conclude that μ<30.

- H_0: $\mu \geq 30$ vs H_A: $\mu < 30$
- Critical value=$Z_{1-\alpha}$=$Z_{0.95}$= 1.645
- Test statistics: $z_{cal} = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \dfrac{27 - 30}{\sqrt{20}/\sqrt{10}} = -2.121$
- Reject H_0 *since* $z = -2.121 < -1.645$
- We conclude that μ is smaller than 30.
- P-value =P(Z < -2.121)=0.0174

Example: Exercise7.2.11 page 236

A random sample 16 emergency reports was selected from files of an ambulance service. The mean time (computed from the sample data) required for ambulance to reach their destinations was 13 minutes. Assume that the population of times is normally distributed with variance of 9. Can we conclude at 0.05 level of significance that the population mean is **greater than 10** minutes?

Solution: Exercise 7.2.11 page 236

- H_0: μ=10 vs H_A: μ>10
- Critical value=$Z_{1-\alpha}$=$Z_{0.95}$= 1.645
- Test statistics: $$Z_{cal} = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{13-10}{\sqrt{9}/\sqrt{16}} = 4.0$$
- Reject H_0 since z = 4.0 >1.645.
- We conclude that μ is greater than 10.
- P-value =P(Z > 4.0) $\longrightarrow$ p-value < 0.0001

Example: Exercise 7.2.17 page 237

Suppose it is known that the IQ scores of a certain population of adults are approximately normally distributed with a standard deviation of 15. A simple random sample of 25 adults drawn from this population had a mean IQ score of 105. On the basis of these data can we conclude that the mean IQ score is **not** **100**? (α=0.05)

Z_{cal}=1.67, p-value=0.095, Don't reject H_0

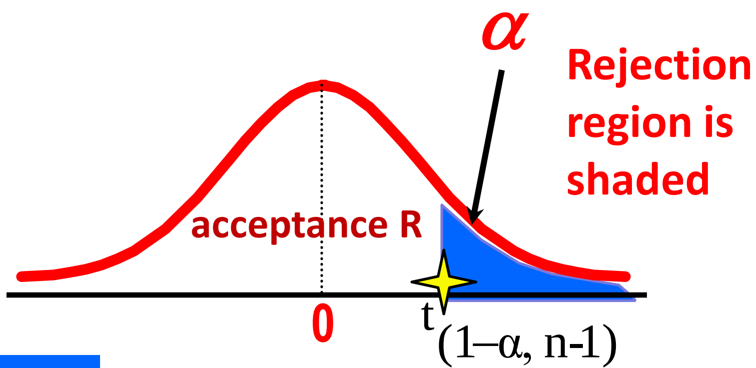
B: Sampling From Normally Distributed population with unknown σ (n small)

Null hypothesis $H_0: \mu = \mu_0$,

α=level of significance

Test statistics: $t_{cal} = \dfrac{\overline{X} - \mu_0}{s/\sqrt{n}}$

a) .



$$\text{Alternative } H_A : \mu > \mu_o$$

$$\text{Reject } H_0 \text{ if } t_{cal} > t_{(1-\alpha,\, n\text{-}1)}$$

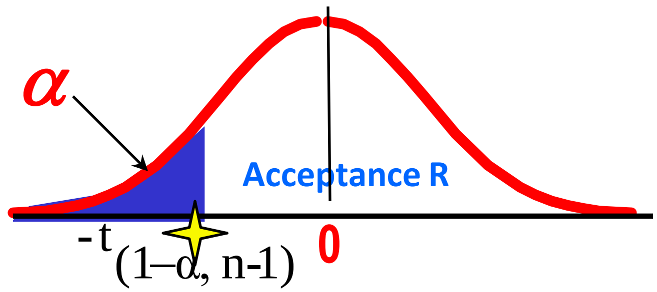
$$P-\text{value} = P\left(t_{n-1} > t_{cal}\right)$$

b)

$$\text{Alternative } H_A : \mu < \mu_o$$

$$\text{Reject } H_0 \text{ if } t_{cal} < -t_{(1-\alpha,\, n-1)}$$

$$P-\text{value} = P\left(t_{n-1} < t_{cal}\right)$$



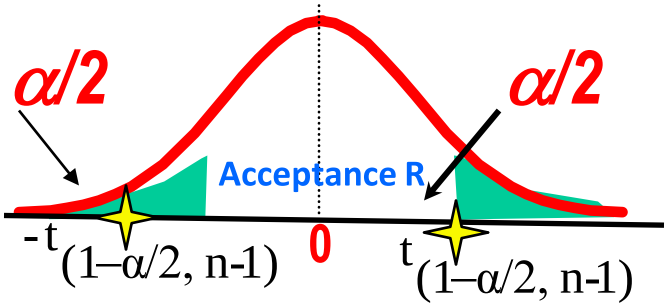
Represents critical value

c)

$$\text{Alternative } H_A : \mu \neq \mu_o$$

$$\text{Reject } H_0 \text{ if } \left|t_{cal}\right| > t_{(1-\alpha/2,\, n-1)}$$

$$P-\text{value} = 2P\left(t_{n-1} > \left|t_{cal}\right|\right)$$



Example: Exercise 7.2.6 page 235

Nine Laboratory animals were infected with a certain bacterium and then immunosuppressed (كبت المناعة). The mean number of organism later recovered from tissue specimens was 6.5 (coded data) with a standard deviation of 0.6. Can we conclude from theses data that the population mean is greater than 6? Let $\alpha=0.05$. What assumption are necessary?

Solution: Exercise 7.2.6 page 235

- Assumption: the population is normally distributed.
- H_0: $\mu \leq 6$ vs H_A: $\mu > 6$
- Critical value= $t_{(1-\alpha,8)} = t_{(0.95,\ 8)} = 1.8595$
- Test statistics: $t_{cal} = \dfrac{\bar{X} - \mu_0}{s/\sqrt{n}} = \dfrac{6.5-6}{0.6/\sqrt{9}} = 2.50$
- Reject H_0 since $t_{cal} = 2.50 > 1.8595$
- P-value = $P(t_8 > 2.50) \longrightarrow 0.01 < \text{p-value} < 0.025$
- we conclude from these data that the population mean is greater than 6

Example: Exercise 7.2.7 page 235

A sample of **25** freshman nursing students made a mean score of **77** on a test designed to measure attitude toward the dying patient. The sample standard deviation was **10**. Do these data provide sufficient evidence to indicate at the **0.05** level of significance, that the population mean is less than 80? What assumptions are necessary?

Solution: Exercise 7.2.6 page 235

- Assumption: the population is normally distributed.
- H_0: $\mu \geq 80$vs H_A: $\mu < 80$
- Critical value= $t_{(1-\alpha,24)}= t_{(0.95,24)}= 1.7109$
- Test statistics: $t_{cal} = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} = \frac{77-80}{10/\sqrt{25}} = -1.5$
- Don't Reject H_0 since $t_{cal} = -1.5 > -1.7109$
- P-value = $P(t_{24} < -1.5) \longrightarrow 0.05 < \text{p-value} < 0.10$
- we conclude from these data that the population mean is not less than 80.

Exercise 7.2.13 Page 236

Can we conclude that the mean maximum voluntary ventilation (التهوية الطوعية) value for apparently health college seniors is not 110 liters per minutes? A sample of 20 yielded the following values:

132, 33, 91, 108, 67, 169, 54, 203, 190, 133, 96, 30, 187, 21, 63, 166, 84, 110, 157, 138

Let α=0.01. What assumptions are necessary?

Mean=111.6 , S=56.303

H_0: μ=110 vs H_A: $\mu \neq$110

t_{cal}=0.1271, p-value>0.2, Don't Reject H_0

C: Sampling From Nonormally Distributed Population (n large)

Null hypothesis H_0: $\mu = \mu_0$, α=level of significance

Test statistics:

$$Z_{cal} = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}, \text{ if } \sigma \text{ known}$$

$$Z_{cal} = \frac{\overline{X} - \mu_0}{s/\sqrt{n}}, \text{ if } \sigma \text{ unknown}$$

a) Alternative $H_A: \mu > \mu_0$, Reject H_0 If $Z_{cal} > Z_{1-\alpha}$

b) Alternative $H_A: \mu < \mu_0$, Reject H_0 If $Z_{cal} < -Z_{1-\alpha}$

c) Alternative $H_A: \mu \neq \mu_0$, Reject H_0 If $|Z_{cal}| > Z_{1-\alpha/2}$

Example: 7.2.4 Page 232

Among 157 African-American men, the mean systolic blood pressure was 146 mmHg with a standard deviation of 27. We wish to know if on the basis of these data, we may conclude that the mean systolic blood pressure for a population of African-American is greater than 140. Use $\alpha=0.05$.

Solution-Example: 7.2.4 Page 232

- H_0: $\mu \leq 140$ vs H_A: $\mu > 140$
- Critical value=$Z_{1-\alpha}$=$Z_{0.95}$= 1.645
- Test statistics: $$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{146 - 140}{27/\sqrt{157}} = 2.78$$
- Reject H_0 *since* $z = 2.78 > 1.645$
- P-value =P(Z > 2.78)= 0.0027
- We conclude that the mean systolic blood pressure for the sampled population is greater than 140.

Example: Exercise 7.2.5 page 235

In a sample of 49 adolescent who served as the subjects in an immunologic study, one variable of interest was the diameter of skin test reaction to an antigen. The sample mean and standard deviation were 21 and 11 mm erythematic, respectively. Can we conclude from these data that the population mean is less than 30? $\alpha=0.05$

- H_0: $\mu \geq 30$ vs H_A: $\mu < 30$.
- Critical value=$Z_{1-\alpha}=Z_{0.95}=1.645$.
- Test statistics: $z=\dfrac{\overline{X}-\mu_0}{s/\sqrt{n}}=\dfrac{21-30}{11/\sqrt{49}}=-5.727$.
- Yes, Reject H_0 *since* $z = -5.727 < -1.645$.
- We conclude that the population mean is not equal to 30.
- P-value =P(Z < -5.727)≈0.000. we write p-value < 0.001.

7.3 HYPOTHESES TESTING: THE DIFFERENCE BETWEEN TWO POPULATION MEANS

A: Sampling From Normally Distributed Populations: Population Variances Known (n small or large)

Null hypothesis $H_0: \mu_1 = \mu_2$ or $H_0: \mu_1 - \mu_2 = 0$, α=l.o.s

Test statistics:

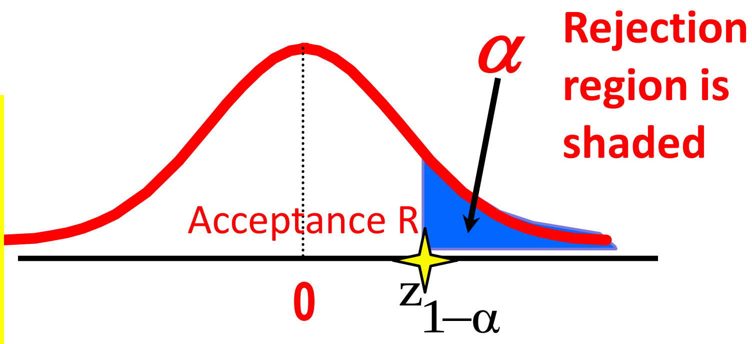
$$Z_{cal} = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

a)

Alternative $H_A: \mu_1 > \mu_2$ or $H_A: \mu_1 - \mu_2 > 0$

Reject H_0 if $Z_{cal} > z_{1-\alpha}$

$P-\text{value} = P\left(Z > Z_{cal}\right)$

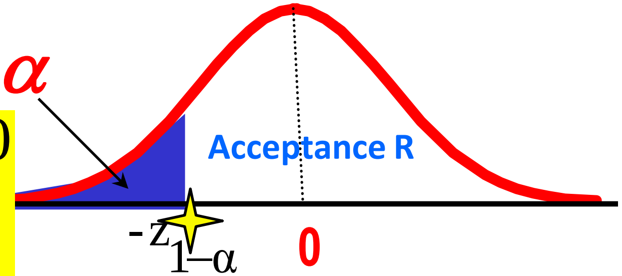


b)

$$\text{Alternative } H_A:\mu_1<\mu_2 \text{ or } H_A:\mu_1-\mu_2<0$$

$$\text{Reject } H_0 \text{ if } Z_{cal}<-z_{1-\alpha}$$

$$P-\text{value}=P\left(Z<Z_{cal}\right)$$

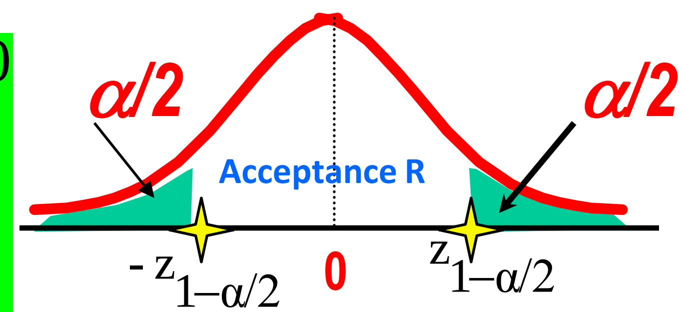


Represents critical value

$$\text{Alternative } H_A:\mu_1\neq\mu_2 \text{ or } H_A:\mu_1-\mu_2\neq0$$

$$\text{Reject } H_0 \text{ if } \left|Z_{cal}\right|>z_{1-\alpha/2}$$

$$P-\text{value}=2P\left(Z>\left|Z_{cal}\right|\right)$$



Example 7.3.1 Page 238

- Researchers wish to know if the data have collected provide sufficient evidence to indicate a difference in mean serum uric acid levels between normal individuals and individual with Down's syndrome. The data consist of serum uric reading on 12 individuals with Down's syndrome from normal distribution with variance 1 and 15 normal individuals from normal distribution with variance 1.5 . The means are $\bar{X}_1 = 4.5\,\mathrm{mg}/100\,\mathrm{ml}$ and $\bar{X}_2 = 3.4\,\mathrm{mg}/100\,\mathrm{ml}$. $\alpha=0.05$.

Solution:

The two populations are normally distributed with known variances.

a) **$H_0: \mu_1 = \mu_2$ vs. $H_A: \mu_1 \neq \mu_2$, $\alpha=0.05$**

b) critical values = $\pm Z_{1-\alpha/2} = \pm Z_{0.975} = \pm 1.96$

c) Test statistic $Z_{cal} = \dfrac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} = \dfrac{4.5-3.4}{\sqrt{\dfrac{1}{12} + \dfrac{1.5}{15}}} = 2.57$

d) Reject H_0 since $Z_{cal} = 2.57 > 1.96$

e) We conclude that there is an indication that the two population means are not equal .

f) P-value $=2P(Z>|Z_{cal}|)=2P(Z>2.57)=2(1-0.9949)=0.0102$

B: Sampling From Normally Distributed Populations
Population Variances Unknown but equal (small samples)

Null hypothesis $H_0: \mu_1 = \mu_2$ or $H_0: \mu_1 - \mu_2 = 0$,
α=level of significance

Test statistics:

$$t_{cal} = \frac{(\bar{X}_1 - \bar{X}_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Where

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

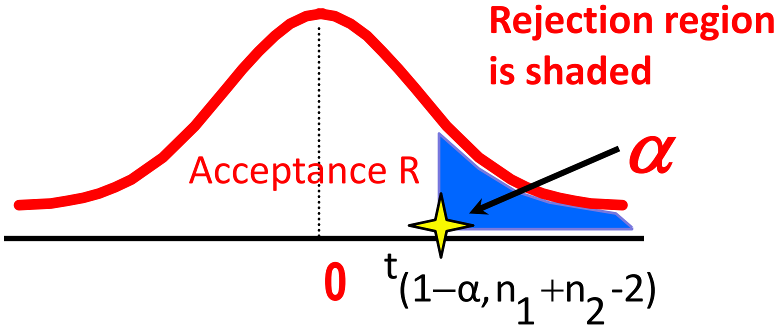
Pooled variance

a)

$$\text{Alternative } H_A: \mu_1 > \mu_2 \text{ or } H_A: \mu_1 - \mu_2 > 0$$

$$\text{Reject } H_0 \text{ if } t_{cal} > t_{(1-\alpha,\, n_1+n_2-2)}$$

$$P\text{-value} = P\left(t_{n_1+n_2-2} > t_{cal}\right)$$

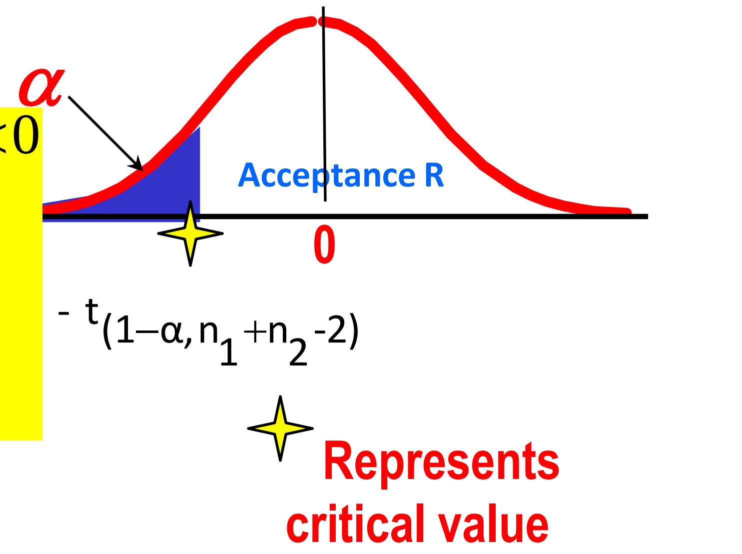


b)

$$\text{Alternative } H_A:\mu_1<\mu_2 \text{ or } H_A:\mu_1-\mu_2<0$$

$$\text{Reject } H_0 \text{ if } t_{cal}<-t_{(1-\alpha,\, n_1+n_2-2)}$$

$$\text{P-value}=P(t_{n_1+n_2-2}<t_{cal})$$

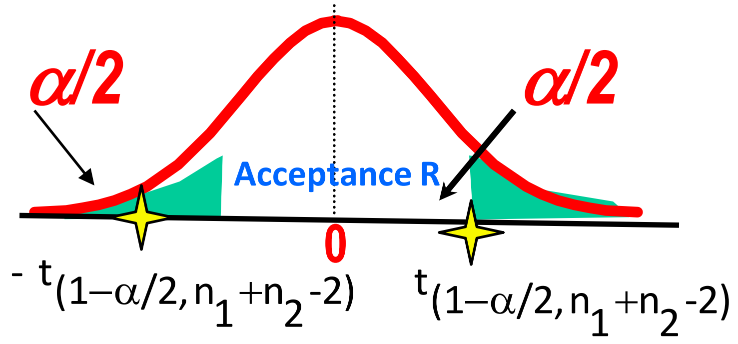


c)

$$\text{Alternative } H_A:\mu_1\neq\mu_2 \text{ or } H_A:\mu_1-\mu_2\neq0$$

$$\text{Reject } H_0 \text{ if } |t_{cal}|>t_{(1-\alpha/2,\, n_1+n_2-2)}$$

$$\text{P-value}=2P(t_{n_1+n_2-2}>|t_{cal}|)$$



Example: Exercise 7.3.6 page 248-249

A test designed to measure mothers' attitude toward their labor and delivery experiences was given to two groups of new mothers. Sample 1 (attenders) had attended prenatal classes held at the local health department. Sample 2 (nonattenders) did not attend the classes. The sample sizes, means, and standard deviations of the test score were as follows:

Sample	n	$\bar{X}$	S
1 (Attenders)	15	4.75	1.0
2 (Nonattenders)	22	3.00	1.5

Does these data provide sufficient evidence to indicate that attenders, on the average, score higher than nonattenders? Let α=0.05.

Solution: Exercise 7.3.6 page 248-249

Assume that the two populations are normally distributed with unknown variances (small samples).

a) $H_0: \mu_1 = \mu_2$ vs. $H_A: \mu_1 > \mu_2$, $\alpha=0.05$

b) critical values = $t_{1-\alpha,35} = t_{0.95,35} = 1.6896$

c) Test statistics:

$$t_{cal} = \frac{(\bar{X}_1 - \bar{X}_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{4.75-3.00}{(1.323).\sqrt{\frac{1}{15}+\frac{1}{22}}} = 3.950$$

$$S_p^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$$

$$= \frac{(15-1).(1)^2+(22-1).(1.5)^2}{15+22-2}$$

$$= 1.75 \Rightarrow S_p = 1.323$$

d) Reject H_0 since $t_{cal} = 3.950 > 1.6896$

e) We conclude that there is an indication that attenders, on the average, score higher than nonattenders .

f) P-value $=P(t_{35}>t_{cal})=P(t_{35} > 3.950)$. P-value < 0.005

Example: Exercise 7.3.7 page 249

Cortisol level determination were made on two samples of women at childbirth. Group 1 subjects underwent emergency cesarean section following induced labor. Group 2 subjects delivered by either cesarean section or the vaginal rout following spontaneous labor. The sample sizes, mean cortisol levels, and standard deviations were as follows:

Sample	n	$\overline{X}$	S
1	10	435	65
2	12	645	80

Does these data provide sufficient evidence to indicate a difference in mean cortisol levels in the populations represented? Let α=0.05.

Solution: Exercise 7.3.7 page 249

Assume that the two populations are normally distributed with unknown variances (small samples).

$$S_p^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2} = \frac{(10-1).(65)^2+(12-1).(80)^2}{10+12-2} = 5421.25 \Rightarrow s_p = 73.629$$

a) $H_0: \mu_1 = \mu_2$ vs. $H_A: \mu_1 \neq \mu_2$, $\alpha=0.05$

b) critical values $= \pm\, t_{1-\alpha/2,20} = \pm\, t_{0.975,20} = \pm\, 2.0860$

c) Test statistics:

$$t_{cal} = \frac{(\bar{X}_1-\bar{X}_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{435-645}{(73.629).\sqrt{\frac{1}{10}+\frac{1}{12}}} = -6.661$$

d) Reject H_0 since $t_{cal} = -6.661 < -2.0860$

e) We conclude that there is an indication that the two population means are not equal .

f) P-value $=2P(t_{20}>|t_{cal}|)=2P(t_{20} >6.661)$. P-value < 0.01

C: Sampling From Nonormally Distributed Population (large samples)

Null hypothesis H_0: $\mu_1 = \mu_2$, α=level of significance

Test statistics:

$$Z_{cal} = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}, \text{ if } \sigma_1^2 \text{ and } \sigma_2^2 \text{ are known}$$

$$Z_{cal} = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}, \text{ if } \sigma_1^2 \text{ and } \sigma_2^2 \text{ are unknown}$$

a) Alternative $H_A: \mu_1 > \mu_2$, Reject H_0 If $Z_{cal} > Z_{1-\alpha}$

b) Alternative $H_A: \mu_1 < \mu_2$, Reject H_0 If $Z_{cal} < -Z_{1-\alpha}$

c) Alternative $H_A: \mu_1 \neq \mu_2$, Reject H_0 If $|Z_{cal}| > Z_{1-\alpha/2}$

Example: 7.3.4 Page 243-244

The objective of a study was to identify the role of various diseases states and additional risk factors in development of thrombosis (الجلطة). One factor of the study was to determine if there were differing levels of the anticariolipin antibody IgG in subjects with and without thrombosis. Table below summarizing the researchers' findings.

Group	Mean IgG level	n	SD
Thrombosis	59.01	53	44.89
No Thrombosis	46.61	54	34.85

We wish to know if we may conclude, on the basis of these results, that in general, persons with thrombosis have, on the average, higher IgG levels than persons without thrombosis. $\alpha=0.01$

Solution: Example 7.3.4 page 243-244

n_1=53, n_2=54, Large samples from independent populations

a) **$H_0: \mu_T < \mu_{NT}$ vs. $H_A: \mu_T \geq \mu_{NT}$, α=0.01**

b) critical values = $Z_{1-\alpha} = Z_{0.99} = 2.33$

c) Test statistics: $$Z_{cal} = \frac{(\bar{X}_T - \bar{X}_{NT})}{\sqrt{\frac{s_T^2}{n_T} + \frac{s_{NT}^2}{n_{NT}}}} = \frac{59.01 - 46.61}{\sqrt{\frac{(44.89)^2}{53} + \frac{(34.85)^2}{54}}} = 1.59$$

d) Don't Reject H_0 since $Z_{cal} = 1.59 < 2.33$

e) These data indicates that on the average, persons with thrombosis and persons without thrombosis may not have different lgG levels

f) P-value $= P(Z > Z_{cal}) = P(Z > 1.59) = 0.0559$

Example: Exercise 7.3.9 page 249

A researcher was interested in knowing if preterm infants with late metabolic acidosis and preterm infants without the condition differ with respect to urine levels of a certain chemical. The mean levels, standard deviation and sample sizes for the two samples studied were as follows:

Sample	n	$\bar{X}$	S
With condition	35	8.5	5.5
Without condition	40	4.8	3.6

What should the researcher conclude on the basis of these results? Let α=0.05.

Solution: Exercise 7.3.9 page 249

The two populations are nonnormally distributed with unknown variances (large samples).

a) H_0: $\mu_1 = \mu_2$ vs. H_A: $\mu_1 \neq \mu_2$, $\alpha=0.05$

b) critical values = $\pm Z_{1-\alpha/2} = \pm Z_{0.975} = \pm 1.96$

c) Test statistics:

$$Z_{cal} = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{8.5 - 3.6}{\sqrt{\frac{(5.5)^2}{35} + \frac{(3.6)^2}{40}}} = 3.394$$

d) Reject H_0 since $Z_{cal} = 3.394 > 1.96$

e) We conclude that there is an indication that the two population means are not equal .

f) P-value $=2P(Z>|Z_{cal}|)=2P(Z > 3.39)=2(0.0003)=0.0006$

Example 7.3.2 Page 240

The purpose of a study was to investigate wheelchair maneuvering in individuals with over-level spinal cord injury (SCI إصابة الحبل الشوكي) and healthy control (C). Subjects used a modified a wheelchair to incorporate a rigid seat surface to facilitate the specified experimental measurements. The data for measurements of the left ischial tuberosity (**الكرسي المتحرك عظام الفخذ وتأثيرها من**) for SCI and control C are shown below

Control	131	115	124	131	122	117	88	114	150	169
SCI	60	150	130	180	163	130	121	119	130	143

We wish to know if we can conclude, on the basis of the above data that the mean of left ischial tuberosity for control C lower than mean of left ischial tuerosity for SCI. Assume both population are normally distributed with unknown variances but assumed equal. $\alpha=0.05$

Solution: Example 7.3.2 page 240

$$n_c = 10, \quad \bar{X}_c = 126.1, \quad s_c = 21.8$$
$$n_{SCI} = 10, \quad \bar{X}_{SCI} = 133.1, \quad s_{SCI} = 32.2$$

$$s_p^2 = \frac{(n_1 - 1)s_c^2 + (n_2 - 1)s_{SCI}^2}{n_1 + n_2 - 2}$$
$$= \frac{(10-1).(21.8)^2 + (10-1).(32.2)^2}{10+10-2}$$
$$= 756.04 \Rightarrow s_p = 27.5$$

a) H_0: $\mu_c \geq \mu_{SCI}$ vs. H_A: $\mu_c < \mu_{SCI}$, $\alpha = 0.05$

b) critical values = $t_{1-\alpha/2,18}$ = $t_{0.95,18}$ = 1.7341

c) Test statistics: $t_{cal} = \dfrac{(\bar{X}_C - \bar{X}_{SCI})}{s_p\sqrt{\dfrac{1}{n_C} + \dfrac{1}{n_{SCI}}}} = \dfrac{126.1 - 133.1}{(27.5).\sqrt{\dfrac{1}{10} + \dfrac{1}{10}}} = -0.569$

d) Don't Reject H_0 since t_{cal} = -0.569 > -1.7341

e) On the basis of these data, we cannot conclude that the population mean pressure is less for healthy subjects than for SCI subjects.

f) P-value $= P(t_{18} > t_{cal}) = P(t_{18} > -0.569)$. P-value > 0.10

7.4 PAIRED COMPARISONS

- In our previous discussion involving the difference between means, it was assumed that the samples were independent.
- A method frequently employed for assessing the effectiveness of a treatment or experimental procedure is one that makes use of related observations resulting from nonindependent (Dependent) samples. A hypothesis test based on this type of data is known as a *paired comparisons test*

- The two-sample t-test compare the means for two groups on a single variable.
- The paired t-test compares the means for two variables for a single group.
- The purpose of this test is to determine whether or not the variables were rated differently by the subjects in the sample.
- This is frequently used to measure change over time, or compare before/after scores, but is can be used more generally to see if one item in a pair was rated higher or lower for the single sample.
- The test actually computes and compares the differences in scores for the two variables, and is equivalent to a one-sample t-test of the computed difference scores.

Data: sample of size n

Subject	Pre (Before)	Post (After)	Difference, $d_i=x_i-y_i$
1	x_1	y_1	d_1
2	x_2	y_2	d_2
3	x_3	y_3	d_3
.	.	.	.
.	.	.	.
.	.	.	.
n	x_n	y_n	d_n

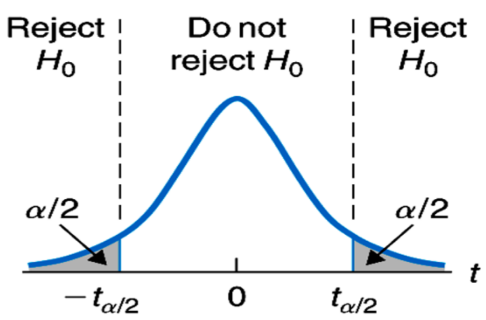
$$\bar{d} = \frac{\sum_{i=1}^{n} d_i}{n}, \qquad s_d = \frac{\sum_{i=1}^{n}(d_i-\bar{d})^2}{n-1} = \frac{\sum_{i=1}^{n} d_i^2-n(\bar{d})^2}{n-1}$$

- **Assumption:** **The observed difference constitute a simple random sample from a normally distributed population**

- **Null hypothesis H_0: $\mu_d = 0$, α=level of significance**

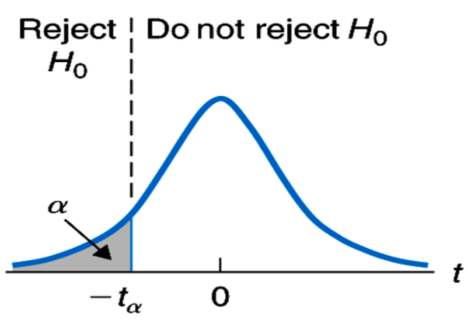
- **Test statistics:** $t_{cal} = \dfrac{\bar{d}}{s_d/\sqrt{n}}$ **, t-distribution with (n-1) *d.f***

H_A: $\mu_d \neq 0$



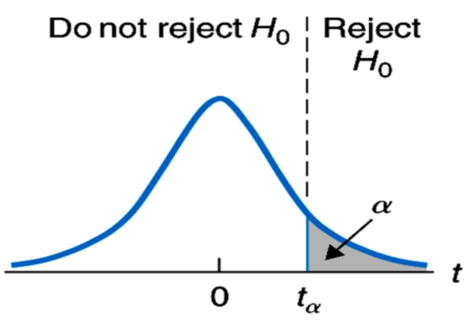
Two-tailed

H_A: $\mu_d < 0$



Left-tailed

H_A: $\mu_d > 0$



Right-tailed

Example:

A physician wants to compare the blood pressures of six patients before and after treatment with a drug. The blood pressures are as follows:

Patient	Before Drug	After Drug
1	168	171
2	171	170
3	182	180
4	167	173
5	174	178
6	170	172

The physician wants to test if there is a significant change of the blood pressure before and after taking the drug at 0.05 level of significance.

Solution:

Subject	Before Drug	After Drug	Difference $d_i=x_i-y_i$	d_i^2
1	168	171	-3	9
2	171	170	1	1
3	182	180	2	4
4	167	173	-6	36
5	174	178	-4	16
6	170	172	-2	4
Total			-12	70

$$\bar{d} = \frac{\sum_{i=1}^{n} d_i}{n} = \frac{-12}{6} = -2$$

$$s_d = \sqrt{\frac{\sum_{i=1}^{n} d_i^2 - n(\bar{d})^2}{n-1}} = \sqrt{\frac{70-6(-2)^2}{5}} = 3.033$$

Solution

- H_0: $\mu_d=0$ vs H_A: $\mu_d \neq 0$, $\alpha=0.05$
- Critical value=$t_{(1-\alpha/2,\ 5)}=t_{(0.975,\ 5)}=$ 2.5706
- Test statistics $t_{cal}=\frac{\bar{d}}{s_d/\sqrt{n}}=\frac{-2}{3.033/\sqrt{6}}=-1.615$
- Don't Reject H_0 *since* $|t_{cal}|= 1.615 < 2.5706$
- P-value =$2P(t_5 > 1.615)$ $\Longrightarrow$ $0.1 < \text{p-value} < 0.2$
- We conclude that there is no significant change of the blood pressure before and after taking the drug

Example:

A group of 10 men were given a special diet for two weeks to test weight loss in pounds. The observed data was:

Man	Weight Before Diet	Weight After Diet
1	181	178
2	171	172
3	190	185
4	187	184
5	210	201
6	202	201
7	166	160
8	173	168
9	183	180
10	184	179

Do the data provide Sufficient evidence to indicate the special diet leads to a weight loss?

α=0.01

Solution:

Man	Weight Before Diet	Weight After Diet	d_i	d_i^2
1	181	178	3	9
2	171	172	-1	1
3	190	185	5	25
4	187	184	3	9
5	210	201	9	81
6	202	201	1	1
7	166	160	6	36
8	173	168	5	25
9	183	180	3	9
10	184	179	5	25
			39	221

$$\bar{d} = \frac{\sum_{i=1}^{n} d_i}{n} = \frac{39}{10} = 3.9$$

$$s_d = \sqrt{\frac{\sum_{i=1}^{n} d_i^2 - n(\bar{d})^2}{n-1}}$$

$$= \sqrt{\frac{221 - 10(3.9)^2}{9}}$$

$$= 2.767$$

Solution

- H_0: $\mu_d \leq 0$ vs H_A: $\mu_d > 0$, $\alpha=0.01$
- Critical value=$t_{(1-\alpha, 9)}=t_{(0.99, 9)}= 2.821$
- Test statistic $t_{cal} = \dfrac{\bar{d}}{s_d/\sqrt{n}} = \dfrac{3.9}{2.767/\sqrt{10}} = 4.46$
- Reject H_0 *since* $t_{cal}= 4.46 > 2.821$
- P-value =$P(t_9 > 4.46) < 0.005$
- Based on the data, we conclude that the special diet leads to a weight loss.