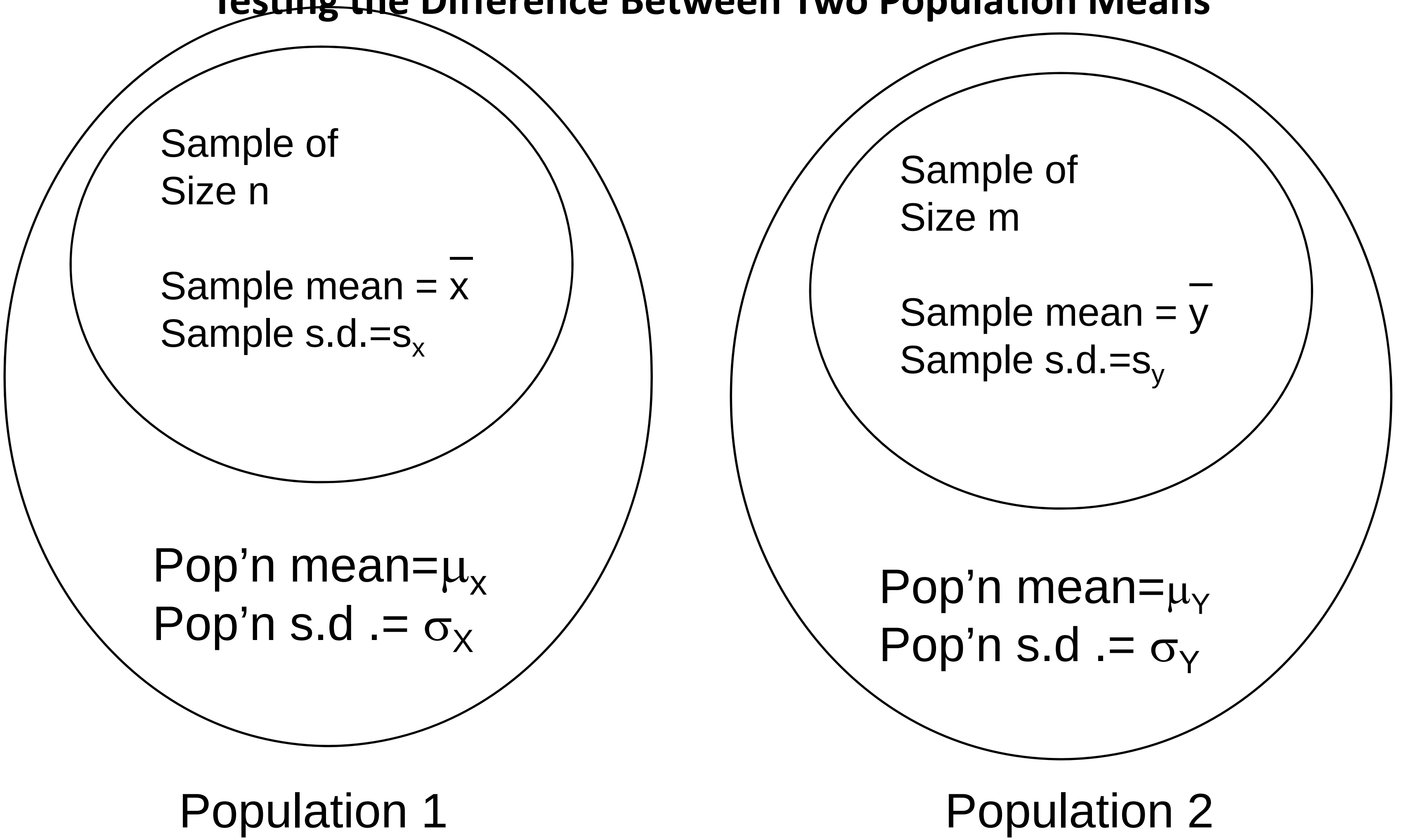
Two Population Case

Testing the Difference Between Two Population Means



Set-up when testing the difference between two population means.

Ho: $\mu_X - \mu_Y = 0$ The means of the 2 populations are the same.

Possible Alternatives (Ha):

$\mu_X - \mu_Y < 0$ The mean of Population 1 is smaller than the mean of Population 2.

$\mu_X - \mu_Y > 0$ The mean of Population 1 is larger than the mean of Population 2.

$\mu_X - \mu_Y \neq 0$ The two population means are different.

Types of Sampling

Two Independent Samples	Two Related Samples or Paired Observations
▪ the selection of the sample in Population 1 is independent of the selection of the sample in Population 2	▪ achieved by either using the same subject in the two samples or pairing of subjects with respect to some extraneous variable that may affect or influence the outcome ▪ used to overcome the difficulty imposed by extraneous differences between the two populations

Testing the Difference Between Two Population Means Based On Two Independent Samples

Ho: $\mu_X - \mu_Y = 0$

Test statistic	Ha	Region of Rejection
Case 1: σ_X & σ_Y known $z_c = \dfrac{\bar{x} - \bar{y}}{\sqrt{\dfrac{\sigma_X^2}{n} + \dfrac{\sigma_Y^2}{m}}}$	$\mu_X - \mu_Y < 0$ $\mu_X - \mu_Y > 0$ $\mu_X - \mu_Y \neq 0$	$z_c < -z_\alpha$ $z_c > z_\alpha$ $z_c < -z_{\alpha/2}$ & $z_c > z_{\alpha/2}$

Case 2: σ_X & σ_Y unknown but equal $$t_c = \frac{\bar{x} - \bar{y}}{\sqrt{S_P^2\left(\frac{1}{n} + \frac{1}{m}\right)}}$$ Where $$S_P^2 = \frac{(n-1)s_X^2 + (m-1)s_Y^2}{n+m-2}$$	$\mu_X - \mu_Y < 0$ $\mu_X - \mu_Y > 0$ $\mu_X - \mu_Y \neq 0$	$t_c < -t_{(\alpha,v)}$ $t_c > t_{(\alpha,v)}$ $t_c < -t_{(\alpha/2,v)}$ & $t_c > t_{(\alpha/2,v)}$ $v = n + m - 2$

Case 3: σ_X & σ_Y unknown & unequal $t_c = \dfrac{\bar{x} - \bar{y}}{\sqrt{\dfrac{s_X^2}{n} + \dfrac{s_Y^2}{m}}}$	$\mu_X - \mu_Y < 0$ $\mu_X - \mu_Y > 0$ $\mu_X - \mu_Y \neq 0$	$t_c < -t_{(\alpha,v)}$ $t_c > t_{(\alpha,v)}$ $t_c < -t_{(\alpha/2,v)}$ & $t_c > t_{(\alpha/2,v)}$ where $v = \dfrac{\left(\dfrac{s_X^2}{n} + \dfrac{s_Y^2}{m}\right)^2}{\dfrac{\left(\dfrac{s_X^2}{n}\right)^2}{n-1} + \dfrac{\left(\dfrac{s_Y^2}{m}\right)^2}{m-1}}$

Testing the Difference Between Two Population Means Based on 2 Independent Samples

- These tests are exact α-level tests for independent samples selected from normal populations. However, they provide good approximate α-level tests when the distributions are not Normal provided both samples are greater than 30.
- **Case 4**: If σ_X and σ_Y are unknown and both sample sizes are greater than 30, use the z-test statistic but replacing σ_X^2 and σ_Y^2 by s_X^2 and s_Y^2, respectively.
- If there is no information at all about the population variances, the test based on the t-test statistic will still provide a good approximate α-level test so long as the sample sizes are the same and both populations are Normal. This is the reason why the researcher should plan the experiment so that n=m.

Example: Testing the Difference Between Two Population Means (independent samples)

A study was conducted to compare the length of time it took male and female students from the same year level and college to answer a 50-item IQ test. Independent samples of 50 male students and 50 female students were asked to take the test in which each person was timed. The results were as follows:

MALE	FEMALE
$n = 50$	$m = 50$
$\bar{x} = 42$ min	$\bar{y} = 38$ min
$s_x^2 = 18$	$s_y^2 = 14$

Did the data present sufficient evidence to suggest a difference between the true mean completion times of male and female students at the 5% level of significance?

Solution:

1) Here we wish to test

   Ho: $\mu_1 = \mu_2$ or equivalently, $\mu_1 - \mu_2 = 0$

   Ha: $\mu_1 \neq \mu_2$ or $\mu_1 - \mu_2 \neq 0$

   μ_1 is the mean length of time it took male students to finish a 50 item IQ test

   μ_2 is the mean length of time it took female students to finish a 50 item IQ test

Con't of Solution:

2) The appropriate test statistic is

$$Z_c = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{S_x^{\ 2}}{n} + \frac{S_y^{\ 2}}{m}}}$$

since both sample sizes are large, the normal distribution applies.

Con't of Solution:

3) Using $\alpha = 0.05$, we find the Critical region:

$$z_c < -z_{\alpha/2} \quad \text{or} \quad z_c > z_{\alpha/2}$$

$$z_c < -z_{0.025} \quad \text{or} \quad z_c > z_{0.025}$$

$$z_c < -1.96 \quad \text{or} \quad z_c > 1.96$$

4) The observed level of the test statistic is

$$Z_c = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}} = \frac{42 - 38}{\sqrt{\frac{18}{50} + \frac{14}{50}}} = 5.0$$

Con’t of Solution:

5) Since 5.0 > 1.96, we thus reject the null hypothesis at the 5% level of significance.

It appears that the difference between completion times of male and female students is significantly different from zero.

Example of Testing the Difference Between Two Population Means (Independent Samples)

During 1985, approximately 31% of Asian women in their late teens and early 20s reported that they had used marijuana within the previous year.

These are the prime reproductive years for women and researchers were concerned about possible effects of marijuana use during pregnancy on fetal growth and development.

A large-scale study of mothers recruited over a three-year period from a general prenatal clinic was done.

Among other findings, the researchers reported that the mean birth weight of infants born to 895 mothers who did not use marijuana during pregnancy was 3260 grams, with a standard deviation of 616 grams.

Con’t of Example

The mean birth weight of infants born to 202 mothers who did use marijuana during pregnancy was 2980 grams, with a standard deviation of 662 grams.

Can the researchers conclude that the birth weights of infants born to women who use marijuana during pregnancy are lower than those of infants born to women who do not? Test at 0.01 level of significance.

PhStat Output

t Test for Differences in Two Means

Data	
Hypothesized Difference	**0**
Level of Significance	**0.01**
Population 1 Sample	
Sample Size	**895**
Sample Mean	**3260**
Sample Standard Deviation	**616**
Population 2 Sample	
Sample Size	**202**
Sample Mean	**2980**
Sample Standard Deviation	**662**

Intermediate Calculations	
Population 1 Sample Degrees of Freedom	894
Population 2 Sample Degrees of Freedom	201
Total Degrees of Freedom	1095
Pooled Variance	390247.2
Difference in Sample Means	280
t-Test Statistic	5.754029

Upper-Tail Test	
Upper Critical Value	**2.329758**
***p*-Value**	**5.65E-09**
Reject the null hypothesis	

Testing the Difference Between two Population Means Based on two Related Samples

Preliminary Steps

1. For each of the n pairs of observations, compute for the difference in scores.

2. Get the mean of the differences between scores, d:

$$\bar{d} = \frac{sum\ of\ \ all\ \ the\ \ differences}{n}$$

3. Compute for the standard deviation of the differences between scores, s_d:

$$\sqrt{\frac{sum\ of\ \ the\ squared\ \ differences - n\bar{d}^2}{n-1}}$$

The Test

Ho: $\mu_X - \mu_Y = 0$

Test statistic: $t = \dfrac{\bar{d}}{\frac{s_d}{\sqrt{n}}}$

Alternative Hypothesis	Region of Rejection (v=n-1)
$\mu_X - \mu_Y < 0$	$t < -t_{\alpha,v}$
$\mu_X - \mu_Y > 0$	$t > t_{\alpha,v}$
$\mu_X - \mu_Y \neq 0$	$t < -t_{\alpha/2},v$ & $t > t_{\alpha/2,v}$

Note: This test is an exact α-level test if the differences between scores come from a normal distribution. However, this still provides a good approximate α-level test even if the distribution is not Normal provided that n>30.

Example: Testing the Difference Between Two Population Means (Paired Samples)

Subjects are tested for reactions times with their left and right hands. (Only right-handed subjects were used.) The results (in thousandths of a second) are given in the accompanying table. Use a 0.05 significance level to test the claim that there is a difference between the mean of the right- and left-hand reaction times. If an engineer is designing a fighter-jet cockpit and must locate the ejection-seat activator to be accessible to either the right or the left hand, does it make a difference which hand she chooses?

Subject	A	B	C	D	E	F	G	H	I	J	K	L	M	N
Right	191	97	116	165	116	129	171	155	112	102	188	158	121	133
Left	224	171	191	207	196	165	177	165	140	188	155	219	177	174

Solution:

1) Ho: $\mu_d = 0$ vs. Ha: $\mu_d \neq 0$ (original claim)

   where

   μ_1 is the mean reaction time of right-hand persons

   μ_2 is the mean reaction time of left-hand persons

2) Because we are testing a claim about the means of paired dependent data, we use the Student t distribution.

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

Con’t of Solution:

3) For α= 0.05 and v = n -1 = 14-1 df, the critical values are t = -2.160 and t = 2.160.

4) Before finding the value of the test statistic, we must first find the values of and . When we evaluate the difference d for each subject, we find these differences d = right – left:

-33	-74	-75	-42	-80	-36	-6
-10	-28	-86	33	-61	-56	-41

Con't of Solution:

where $\bar{d} = \frac{\sum d}{n} = \frac{-595}{14} = -42.5$

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

$$= \sqrt{\frac{14(39{,}593) - (-595)^2}{14(14-1)}}$$

$= 33.2$

Con’t of Solution:

With these statistics and the assumption that

μ_d = 0, we can now find the value of the test statistic:

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

$$= \frac{-42.5 - 0}{33.2 / \sqrt{14}} = -4.790$$

Con't of Solution:

5) Because the test statistic does fall in the critical region, we reject the null hypothesis of $\mu_d = 0$. There is sufficient evidence to support the claim of a difference between the right- and left-hand reaction times.

Because there does appear to be such a difference, an engineer designing a fighter-jet cockpit should locate the ejection-seat activator so that it is readily accessible to the faster hand, which appears to be the right hand with seemingly lower reaction times. (We could require special training for left-handed pilots if a similar test of left-handed pilots show that their dominant hand is faster.)

Example: Testing the Difference Between Two Population Means (Paired Samples)

A public school is considering the revision of its Reading course.

The school believes that the revised course will improve the reading comprehension skills of its students.

The school has decided to conduct an experiment to evaluate the revised course before it is offered to the general student body.

Pairs of students were formed based on their grades in the Reading course the previous semester.

A random sample of 10 pairs was then selected. A randomization process was again used to determine which student in the pair would be taught using the revised course.

At the end of the course, all the students in the sample were given the same exam. Based on the scores below, can it be concluded at α=0.05 level of significance that the revised course has improved the reading comprehension of the students?

	Reading Exam Scores	
Pair	Existing Course	Revised Course
1	211	221
2	231	216
3	191	203
4	216	224
5	207	201
6	203	178
7	201	188
8	179	159
9	179	177
10	211	197

t-Test: Paired Two Sample for Means

Output from EXCEL

	Existing Course	*Revised Course*
Mean	202.9	196.4
Variance	266.3222222	444.4888889
Observations	10	10
Pearson Correlation	0.78423019	
Hypothesized Mean Difference	0	
df	9	
t Stat	1.571099752	
P(T<=t) one-tail	0.075303908	
t Critical one-tail	1.833113856	
P(T<=t) two-tail	0.150607816	
t Critical two-tail	2.262158887	

Testing the Difference Between Two Population Proportions

When testing a hypothesis made about two population proportions – such as proportions of cured patients in a population given some treatment and a second population given a placebo – we make the following assumptions and use the following notation.

<u>Assumptions:</u>

- We have two independent sets of randomly selected sample data.
- For both samples, the conditions $np \geq 5$ and $nq \geq 5$ are satisfied.

Notations for Two Proportions

For population 1 we let

p_1 = population proportion

n_1 = size of the sample

x_1 = number of successes in the sample

$$\hat{p}_1 = \frac{x_1}{n_1}$$ (the sample proportion)

$$\hat{q}_1 = 1 - \hat{p}_1$$

The corresponding meanings are attached to p_2, n_2, x_2, $\hat{p}_2$, and $\hat{q}_2$, which come from population 2.

Test Statistic for Two Proportions

The following test statistic applies to null and alternative hypotheses that fit one of these three formats:

Ho: $p_1 = p_2$ Ho: $p_1 = p_2$ Ho: $p_1 = p_2$

Ha: $p_1 \neq p_2$ Ha: $p_1 < p_2$ Ha: $p_1 > p_2$

$$Z_c = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\hat{p}\hat{q}}{n_1} + \frac{\hat{p}\hat{q}}{n_2}}}$$

where $p_1 - p_2 = 0$

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} \qquad \hat{q} = 1 - \hat{p} \qquad \hat{p}_1 = \frac{x_1}{n_1} \qquad \hat{p}_2 = \frac{x_2}{n_2}$$

Example of Testing the Difference Between Two Proportions:

- Johns Hopkins researchers conducted a study of pregnant IBM employees.

  Among 30 employees who worked with glycol ethers, 10 (or 33.3%) had miscarriages, but among 750 who were not exposed to glycol ethers, 120 (or 16.0%) had miscarriages.

  At the 0.01 significance level, test the claim that the miscarriage rate is greater for women exposed to glycol ethers.

Solution:

We stipulate that sample 1 is the group that worked with glycol ethers and sample 2 is the group not exposed, so the sample statistics can be summarized as shown here:

Exposed to Glycol Ethers	Not Exposed to Glycol Ethers
$n_1 = 30$	$n_2 = 750$
$x_1 = 10$	$x_2 = 120$
$\hat{p}_1 = 10/30 = 0.333$	$\hat{p}_2 = 120/750 = 0.160$

Con’t of Solution:

1) Ho: $p_1 = p_2$ vs. Ha: $p_1 > p2$ (original claim)
2) We will use the normal distribution as an approximation to the binomial distribution. We have two independent samples, and the conditions $np \geq 5$ and $nq \geq 5$ are satisfied for each of the two samples. To check this, we note that in conducting this test, we assume that $p_1 = p_2$, where their common value is the pooled estimate , calculated as

- $\hat{p} = \dfrac{x_1 + x_2}{n_1 + n_2}$ = $\dfrac{10 + 120}{30 + 750} = 0.1667$

Con’t of Solution:

With $\hat{p} = 0.1667$,

it follows that $\hat{q} = 1 - \hat{p}$

$= 1 - 0.1667 = 0.8333.$

We verify that np $\geq$ 5 and nq $\geq$ 5 for both samples as follows:

Sample 1	**Sample 2**
$n_1 p$ = (30)(0.1667) = 5	$n_2 p$ =(750)(0.1667) = 12.5
$n_1 q$ = (30)(0.8333) = 25	$n_2 q$ = (750)(0.8333) = 62.5

Con’t of Solution:

3) The critical value is $z = 2.33$ with 0.01 level of significance.

4) The test statistic is:

$$Z_c = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\hat{p}\hat{q}}{n_1} + \frac{\hat{p}\hat{q}}{n_2}}}$$

$$z_c = \frac{(0.333 - 0.160) - 0}{\sqrt{\frac{(0.1667)(0.8333)}{30} + \frac{(0.1667)(0.8333)}{750}}} = 2.49$$

Con’t of Solution:

5) The test statistic falls within the critical region, so we reject the null hypothesis of p1 = p2.

   We conclude that there is sufficient evidence to support the claim that the miscarriage rate is greater for women exposed to ethyl glycol.

- With this evidence, the John Hopkins researchers concluded that women employees exposed to glycol ethers “have a significantly increased risk of miscarriage.” On the basis on these results, IBM warned its employees of the danger, notified the Environmental Protection Agency, and greatly reduced its use of glycol ethers.

Example: Testing the Difference Between Two Proportions

- The newly appointed head of a mental health agency claims that a greater proportion of the crimes committed by persons younger than 21 years of age are violent crimes.

  Of 2750 randomly selected arrests of criminals younger than 21 years of age, 4.25% involve violent crimes.

  Of 2200 randomly selected arrests of criminals 21 years of age or older, 4.55% involve violent crimes.

  Does this indicate that there isn't a significant difference between the two rates of violent crimes? Use 0.05 level of significance.

PhStat Output

Z Test for Differences in Two Proportions

Data	
Hypothesized Difference	**0**
Level of Significance	**0.05**
Group 1	
Number of Successes	**117**
Sample Size	**2750**
Group 2	
Number of Successes	**100**
Sample Size	**2200**
Intermediate Calculations	
Group 1 Proportion	0.042545455
Group 2 Proportion	0.045454545
Difference in Two Proportions	-0.002909091
Average Proportion	0.043838384
Z Test Statistic	-0.496751813
Two-Tailed Test	
Lower Critical Value	**-1.959962787**
Upper Critical Value	**1.959962787**
***p*-Value**	**0.619364071**
Do not reject the null hypothesis	