Chapter 9. Comparing Two Population Means

9.1 Introduction

Two Population Distributions

- In Chapter 8, we considered a single population distribution
  - For example, we estimated the mean of a single population distribution
- But it is important to make a comparison between two population distributions

Two Population Distributions

- Two-sample problem:

  Set of data observations from population A

  $$x_1, x_2, x_3, \dots, x_n$$

  Additional set of observations from population B

  $$y_1, y_2, y_3, \dots, y_m$$

- Sample of data observations x_i are independent observations from the unknown probability distribution of A
- Sample of data observations y_i are independent observations from the unknown probability distribution of B
- Sample sizes *n* and *m* need not be equal (but experiments are often designed to have equal sample sizes)

Two Population Distributions

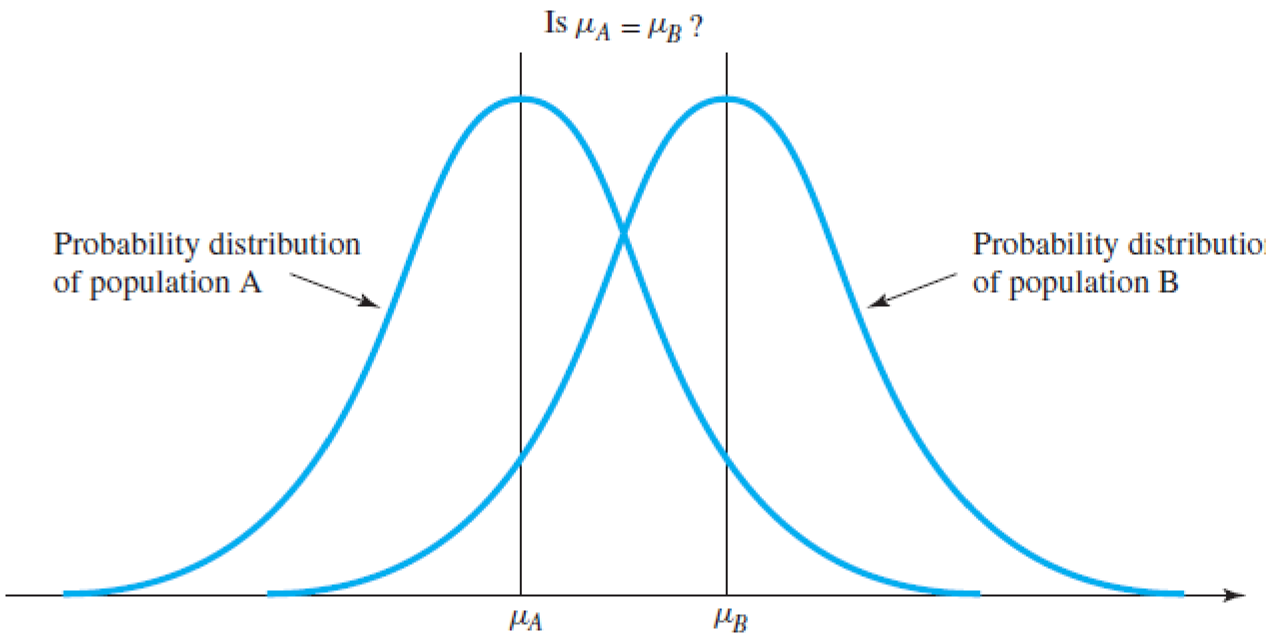
- *Example*:

  Population 1: Height of students in Korea

  Population 2: Height of students in Europe

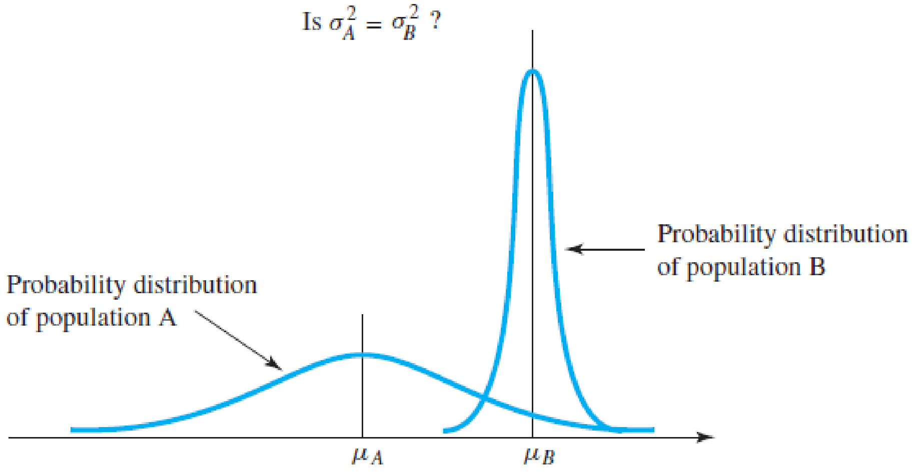
Two Population Distributions

- In general, interested in assessing evidence that there is a difference between the two probability distributions
- One approach is to compare the means of the two probability distributions



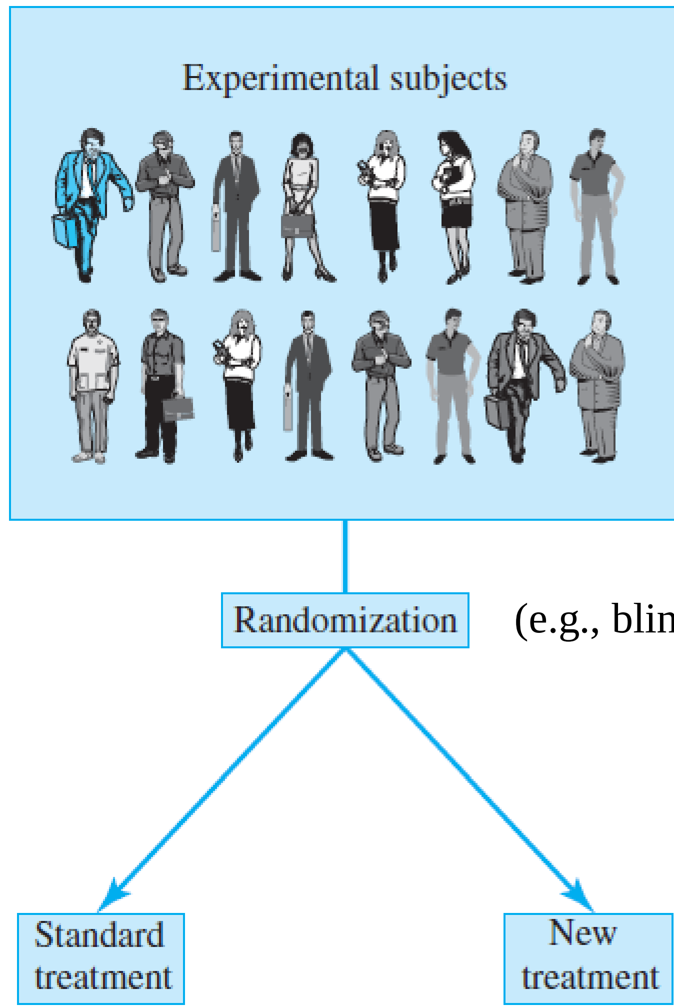
Two Population Means

- If we find that $\mu_A \neq \mu_B$, then this indicates that the two probability distributions are different
- If we find evidence that $\mu_A = \mu_B$,
  - we may conclude that the two probability distributions may be identical
  - or we may further compare the variances of the two data sets



Two Population Means

- *Example*:



(e.g., blind, double-blind)

Control group

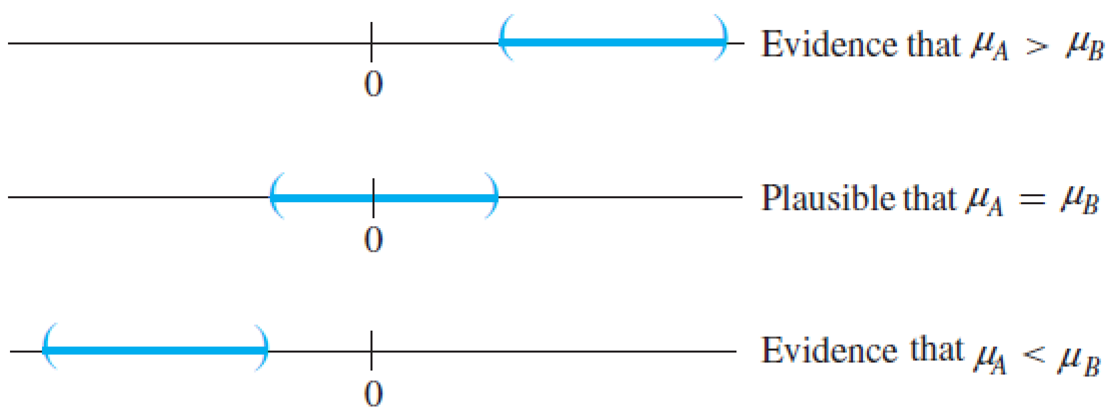
Two Population Means

- How do we compare μ_A and μ_B?

  ➔ Since we want to see if the two are the same, we construct a confidence interval for

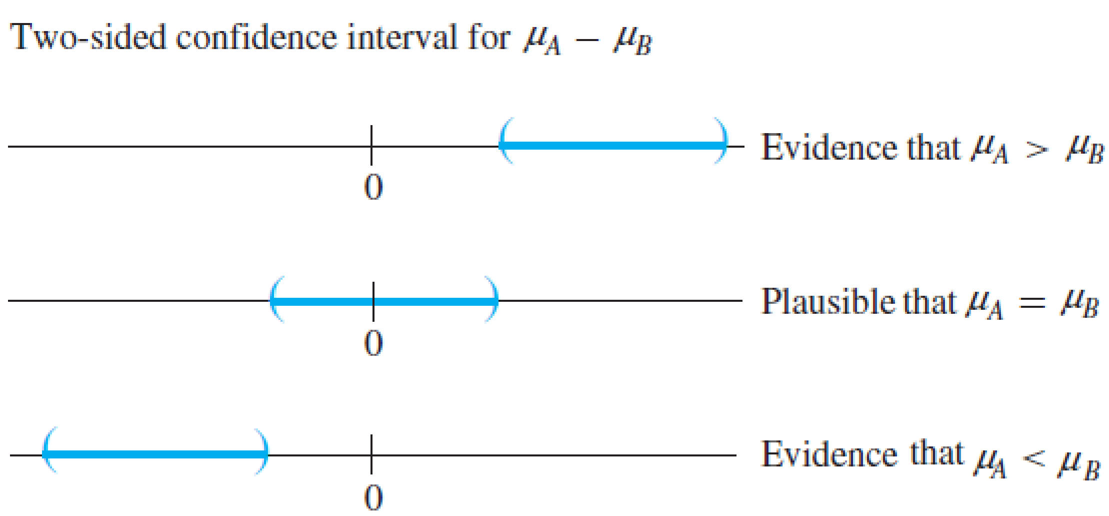
$$\mu_A - \mu_B$$

Two-sided confidence interval for $\mu_A - \mu_B$



Two Population Means

- How do we compare μ_A and μ_B?
  - ➔ Since we want to see if the two are the same, we construct a confidence interval for $\mu_A - \mu_B$
- We are interested whether this confidence interval contains zero



Two Population Means

- How do we compare μ_A and μ_B?
  - ➔ Since we want to see if the two are the same, we construct a confidence interval for $\mu_A - \mu_B$
- We are interested whether this confidence interval contains zero
- The confidence interval is centered at $\bar{x} - \bar{y}$ (from our samples)

Two Population Means

- How do we compare μ_A and μ_B?

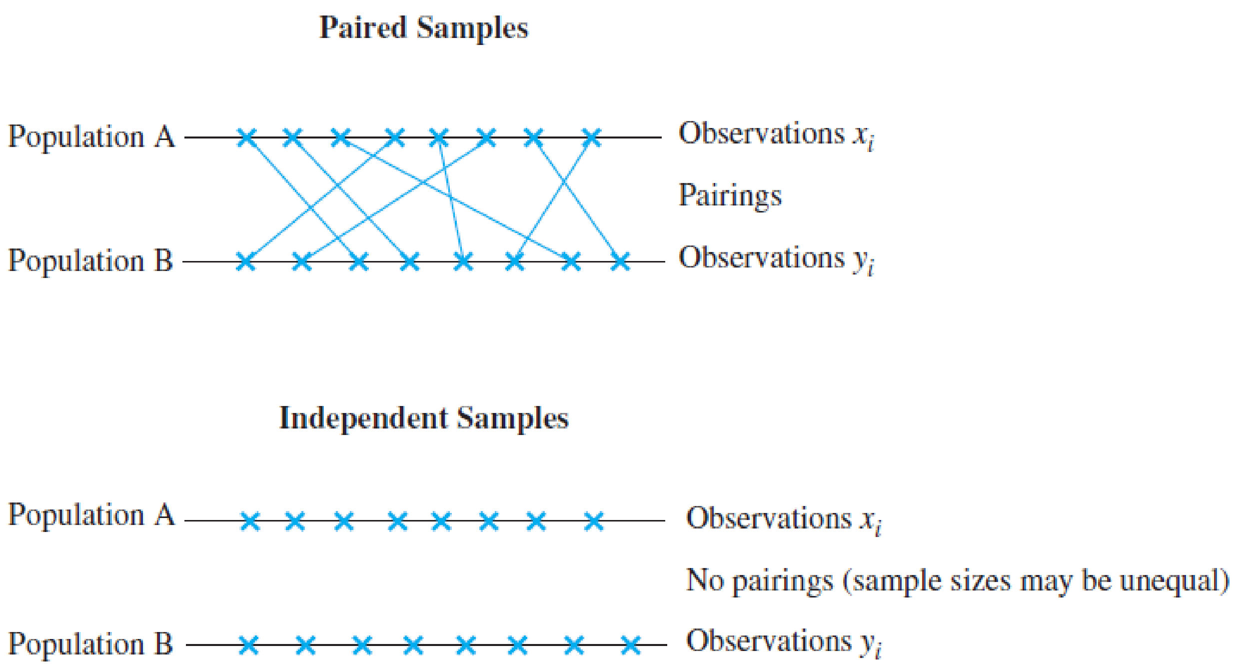
  ➔ Another approach is to perform a hypothesis test

$$H_0 : \mu_A = \mu_B \quad \text{versus} \quad H_A : \mu_A \neq \mu_B$$

- Then, a small *p*-value will indicate that the null hypothesis is not plausible

Paired vs. Independent Samples

- Two data sets may be paired samples or independent samples
- Paired samples may alleviate variability from outside factors



Paired Samples

- **Paired samples** can be expressed as

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$$

- Samples sizes from the two populations must be equal
- Comparison between the two is then based upon the pairwise differences:

$$z_i = x_i - y_i \qquad 1 \le i \le n$$

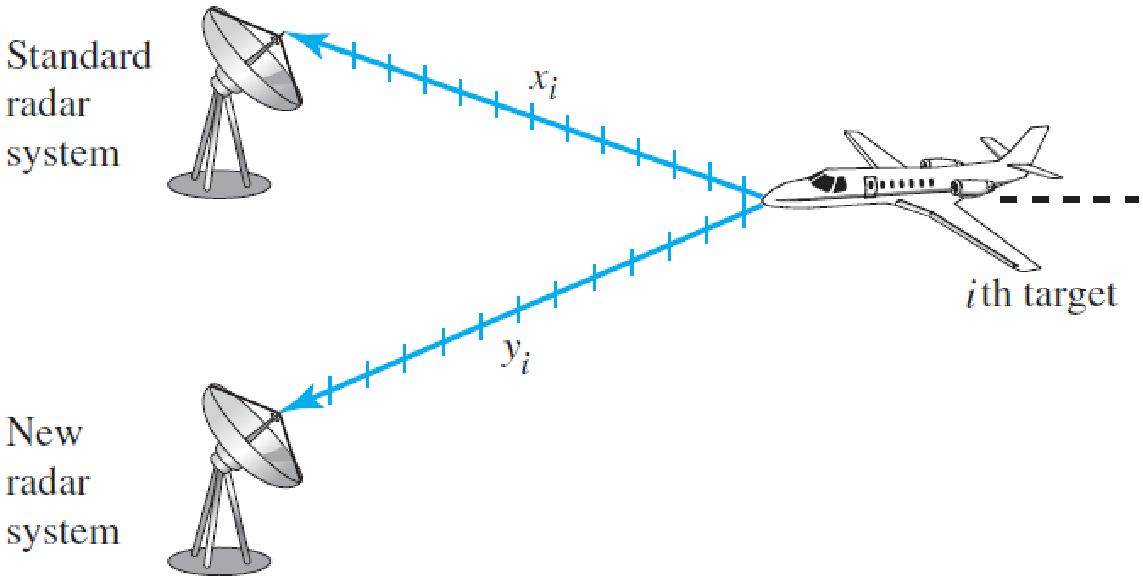
Paired Samples

- *Example*:

	Day 1	Day 2	Difference
patient 1	standard drug x_1	new drug y_1	$z_1 = x_1 - y_1$
patient 2	new drug y_2	standard drug x_2	$z_2 = x_2 - y_2$
patient 3	standard drug x_3	new drug y_3	$z_3 = x_3 - y_3$
patient 4	new drug y_4	standard drug x_4	$z_4 = x_4 - y_4$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
patient 39	standard drug x_{39}	new drug y_{39}	$z_{39} = x_{39} - y_{39}$
patient 40	new drug y_{40}	standard drug x_{40}	$z_{40} = x_{40} - y_{40}$

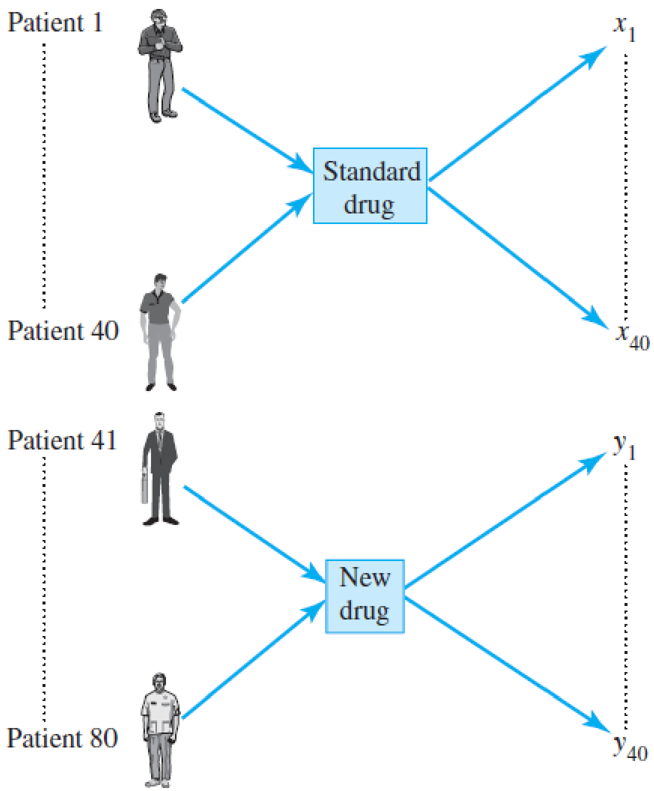
Paired Samples

- *Example*:



Independent Samples

- *Example*:



What if the second set of patients happen to be a group more receptive to drugs?

9.2 Analysis of Paired Samples

Paired Samples

- Analysis of paired samples $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$

  is performed by reducing the problem to a one-sample problem

  ➔ Calculate the differences

$$z_i = x_i - y_i \qquad 1 \leq i \leq n$$

- Data observations z_i can be taken to be independent identically distributed observations from some probability distribution with mean μ
- Then, one-sample techniques from Chapter 8 can be applied to

$$z_1, \ldots, z_n$$

Paired Samples

- Then, the parameter μ is the average difference between A and B

$$\mu = \mu_A - \mu_B$$

  - Positive μ indicates that the mean population of A is larger than the mean population of B
  - Negative μ indicates that the mean population of A is smaller than the mean population of B

- Often valuable to test:

$$H_0 : \mu = 0 \quad \text{versus} \quad H_A : \mu \neq 0$$

➔ If the null hypothesis is plausible, then there is no sufficient evidence that the means of A and B are different

Paired Samples

- Suppose each observation obtained from population A is thought of as below

$$x_i = \mu_A + \gamma_i + \epsilon_i^A$$

  - μ_A: effect of population A
  - γ_i: effect of subject *i*
  - ϵ_i^A: random error (with expectation = 0)

Paired Samples

- Suppose each observation obtained from population A is thought of as below

$$x_i = \mu_A + \gamma_i + \epsilon_i^A$$

- Similarly, suppose each observation obtained from population B is thought of as below

$$y_i = \mu_B + \gamma_i + \epsilon_i^B$$

  - μ_B: effect of population B
  - γ_i: effect of subject *i*
  - ϵ_i^B: random error (with expectation = 0)

Paired Samples

$$x_i = \mu_A + \gamma_i + \epsilon_i^A$$

$$y_i = \mu_B + \gamma_i + \epsilon_i^B$$

- Then, the difference becomes

$$z_i = \mu_A - \mu_B + \epsilon_i^{AB}$$

where the error term is

$$\epsilon_i^{AB} = \epsilon_i^A - \epsilon_i^B$$

Paired Samples

- Since the error term is an observation from a distribution with a zero expectation, the differences z_i are consequently observations from a distribution with expectation

$$\mu = \mu_A - \mu_B$$

which does not depend on the subject effect γ_i

Paired Samples

- *Example*: (Heart rate reduction from new drug)

$$H_0 : \mu = 0 \quad \text{versus} \quad H_A : \mu \neq 0$$

Patient	Standard drug x_i	New drug y_i	$z_i = x_i - y_i$
1	28.5	34.8	−6.3
2	26.6	37.3	−10.7
3	28.6	31.3	−2.7
4	22.1	24.4	−2.3
5	32.4	39.5	−7.1
6	33.2	34.0	−0.8
7	32.9	33.4	−0.5
8	27.9	27.4	0.5
9	26.8	35.4	−8.6
10	30.7	35.7	−5.0
11	39.6	40.4	−0.8
12	34.9	41.6	−6.7
13	31.1	30.8	0.3
14	21.6	30.5	−8.9
15	40.2	40.7	−0.5
16	38.9	39.9	−1.0
17	31.6	30.2	1.4
18	36.0	34.5	1.5
19	25.4	31.2	−5.8
20	35.6	35.5	0.1

Patient	Standard drug x_i	New drug y_i	$z_i = x_i - y_i$
21	27.0	25.3	1.7
22	33.1	34.5	−1.4
23	28.7	30.9	−2.2
24	33.7	31.9	1.8
25	33.7	36.9	−3.2
26	34.3	27.8	6.5
27	32.6	35.7	−3.1
28	34.5	38.4	−3.9
29	32.9	36.7	−3.8
30	29.3	36.3	−7.0
31	35.2	38.1	−2.9
32	29.8	32.1	−2.3
33	26.1	29.1	−3.0
34	25.6	33.5	−7.9
35	27.6	28.7	−1.1
36	25.1	31.4	−6.3
37	23.7	22.4	1.3
38	36.3	43.7	−7.4
39	33.4	30.8	2.6
40	40.1	40.8	−0.7

Heart rate reductions data set
(% reduction in heart rate)

Paired Samples

- *Example*: (Heart rate reduction from new drug)

  From the collected data: $\bar{z} = -2.655,\ s_z = 3.730,\ \ n = 40$

  Then, with $H_0 : \mu = 0$

$$t = \frac{\sqrt{n}(\bar{z} - \mu)}{s} = \frac{\sqrt{40} \times (-2.655)}{3.730} = -4.50$$

  and

$$p\text{-value} = 2 \times P(X > 4.50) \simeq 0.0001$$

  where the random variable X has a t-distribution with 39 degrees of freedom

This analysis reveals that it is *not* plausible that $\mu = 0$, and so the experimenter can conclude that there is evidence that the new drug has a different effect from the standard drug.

Paired Samples

- *Example*: (Heart rate reduction from new drug)

From the critical point $t_{0.005,39} = 2.7079$, a 99% two-sided confidence interval for the difference between the average effects of the drugs is

$$\mu = \mu_A - \mu_B \in \left(\bar{z} - \frac{t_{0.005,39}s}{\sqrt{40}}, \bar{z} + \frac{t_{0.005,39}s}{\sqrt{40}} \right)$$
$$= \left(-2.655 - \frac{2.7079 \times 3.730}{\sqrt{40}}, -2.655 + \frac{2.7079 \times 3.730}{\sqrt{40}} \right)$$
$$= (-4.252, -1.058)$$

Consequently, based upon this data set the experimenter can conclude that the new drug provides a reduction in a patient's heart rate of somewhere between 1% and 4.25% more on average than the standard drug.

Paired Samples

- *Example*: (Radar detection systems)

$$H_0 : \mu \geq 0 \quad \text{versus} \quad H_A : \mu < 0$$

Target	Standard radar system x_i	New radar system y_i	$z_i = x_i - y_i$
1	48.40	51.14	−2.74
2	47.73	46.48	1.25
3	51.30	50.90	0.40
4	50.49	49.82	0.67
5	47.06	47.99	−0.93
6	53.02	53.20	−0.18
7	48.96	46.76	2.20
8	52.03	54.44	−2.41
9	51.09	49.85	1.24
10	47.35	47.45	−0.10
11	50.15	50.66	−0.51
12	46.59	47.92	−1.33
13	52.03	52.37	−0.34
14	51.96	52.90	−0.94
15	49.15	50.67	−1.52
16	48.12	49.50	−1.38
17	51.97	51.29	0.68
18	53.24	51.60	1.64
19	55.87	54.48	1.39
20	45.60	45.62	−0.02
21	51.80	52.24	−0.44
22	47.64	47.33	0.31
23	49.90	51.13	−1.23
24	55.89	57.86	−1.97

Radar detection systems data set (distance of target in miles when detected)

Paired Samples

- *Example*: (Radar detection systems)

From the collected data: $\bar{z} = -0.261, \; s_z = 1.305, \; n = 24$

Then, with $H_0 : \mu \geq 0$

$$t = \frac{\sqrt{n}(\bar{z} - \mu)}{s} = \frac{\sqrt{24} \times (-0.261)}{1.305} = -0.980$$

and

$$P(X \leq -0.980) = 0.170$$

where the random variable X has a t-distribution with 23 degrees of freedom

Large p-value ➔ This data set does not provide sufficient evidence to establish that the new radar system is better than the standard system

9.3 Analysis of Independent Samples

Independent Samples

- **Independent (unpaired) samples**
  - n observations x_i from population A
  - m observations y_i from population B
- Goal: Inference on the difference between population means, $\mu_A - \mu_B$
- Point estimate: $\bar{x} - \bar{y}$

	Sample size	Sample mean	Sample standard deviation
Population A	n	$\bar{x}$	s_x
Population B	m	$\bar{y}$	s_y

Independent Samples

- Point estimate: $\bar{x} - \bar{y}$
- For confidence intervals, we need the point estimate and also the standard error
- What is the standard error of $\bar{x} - \bar{y}$?

Independent Samples

- Point estimate: $\bar{x} - \bar{y}$
- What about the standard error of this estimate?
  - Since $\text{Var}(\bar{x}) = \sigma_A^2/n$ and $\text{Var}(\bar{y}) = \sigma_B^2/m$,

    where σ_A^2 and σ_B^2 are the two population variances, this

    point estimate has a standard error

$$\text{s.e.}(\bar{x} - \bar{y}) = \sqrt{\frac{\sigma_A^2}{n} + \frac{\sigma_B^2}{m}}$$

Independent Samples

- But this includes the population variances of A and B

$$\text{s.e.}(\bar{x} - \bar{y}) = \sqrt{\frac{\sigma_A^2}{n} + \frac{\sigma_B^2}{m}}$$

- What can we do?
  - Use the sample variances of A and B:

$$\text{s.e.}(\bar{x} - \bar{y}) = \sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}$$

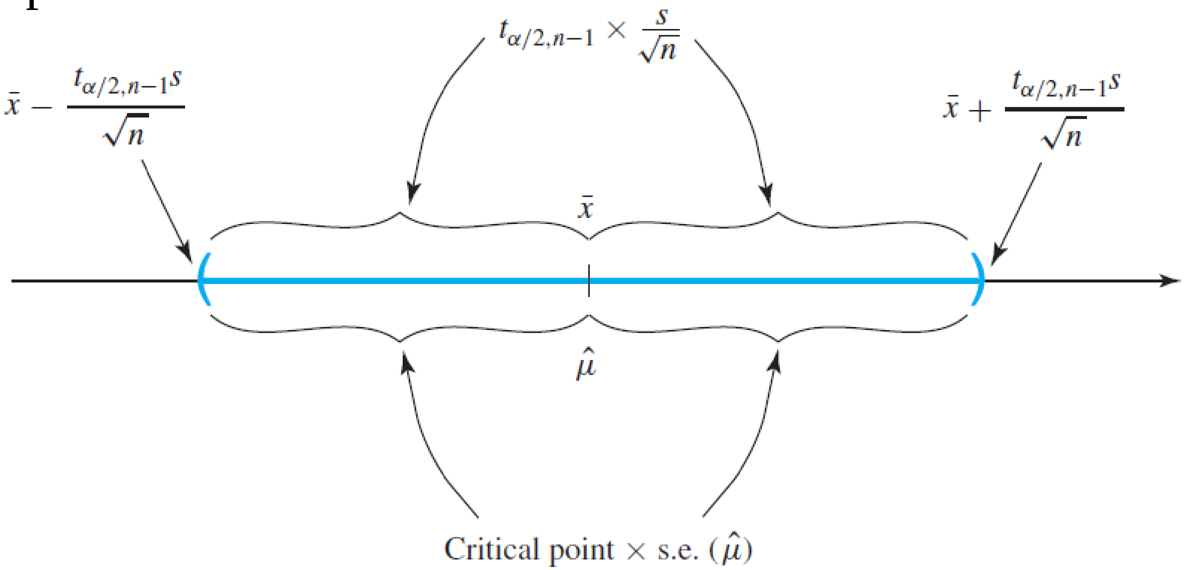
This is the general approach for independent samples

Independent Samples: *t*-Intervals

- *Recall*: The basic structure of *t*-intervals

$$\mu \in (\hat{\mu} - \text{critical point} \times \text{s.e.}(\hat{\mu}),\ \hat{\mu} + \text{critical point} \times \text{s.e.}(\hat{\mu}))$$

- For example:



Independent Samples: *t*-Intervals

$$\mu \in (\hat{\mu} - \text{critical point} \times \text{s.e.}(\hat{\mu}), \hat{\mu} + \text{critical point} \times \text{s.e.}(\hat{\mu}))$$

where $\mu = \mu_A - \mu_B$

- So for independent samples, we need to find the following values:
  - $\hat{\mu}$
  - $\text{s.e.}(\hat{\mu})$
  - critical point

Independent Samples: *t*-Intervals

$$\mu \in (\hat{\mu} - \text{critical point} \times \text{s.e.}(\hat{\mu}), \hat{\mu} + \text{critical point} \times \text{s.e.}(\hat{\mu}))$$

where $\mu = \mu_A - \mu_B$

- Since our goal is the difference $\mu_A - \mu_B$
  - $\hat{\mu} \rightarrow \bar{x} - \bar{y}$
  - $\text{s.e.}(\hat{\mu}) \rightarrow \text{s.e.}(\bar{x} - \bar{y}) = \sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}$
  - critical point: ?

Independent Samples: *t*-Intervals

- It turns out that we can still use the *t*-distribution
- But the degree of freedom is calculated as:

$$\nu = \frac{\left(\frac{s_x^2}{n} + \frac{s_y^2}{m}\right)^2}{\frac{s_x^4}{n^2(n-1)} + \frac{s_y^4}{m^2(m-1)}}$$

***Round down* to the nearest integer**

Independent Samples: *t*-Intervals

$$\mu \in (\hat{\mu} - \text{critical point} \times \text{s.e.}(\hat{\mu}), \hat{\mu} + \text{critical point} \times \text{s.e.}(\hat{\mu}))$$

where $\mu = \mu_A - \mu_B$

- Since our goal is the difference $\mu_A - \mu_B$
  - $\hat{\mu} \rightarrow \bar{x} - \bar{y}$
  - $\mathrm{s.e.}(\hat{\mu}) \rightarrow \mathrm{s.e.}(\bar{x} - \bar{y}) = \sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}$
  - critical point: $t_{\alpha/2,v}$

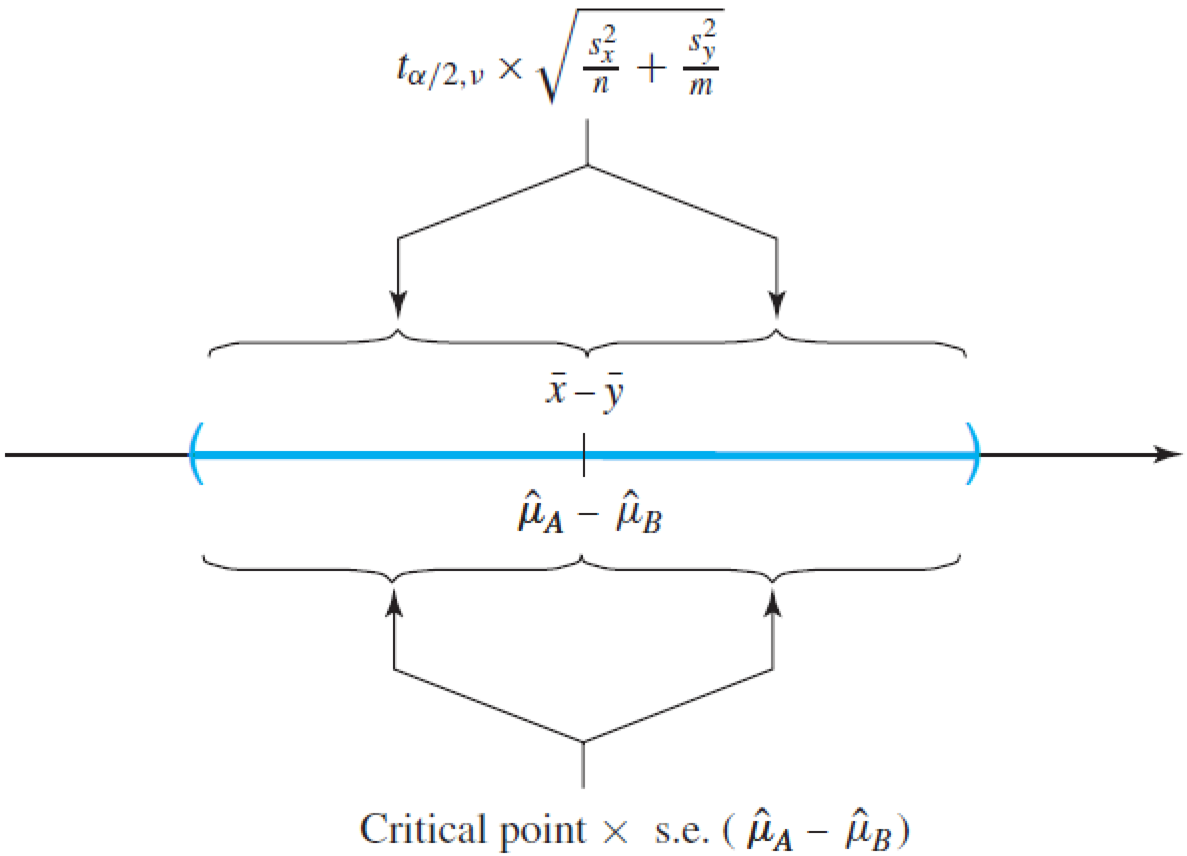
Independent Samples: *t*-Intervals

$$\mu \in (\hat{\mu} - \text{critical point} \times \text{s.e.}(\hat{\mu}),\ \hat{\mu} + \text{critical point} \times \text{s.e.}(\hat{\mu}))$$

A two-sided $1 - \alpha$ level confidence interval for $\mu_A - \mu_B$ is therefore

$$\mu_A - \mu_B \in \left(\bar{x} - \bar{y} - t_{\alpha/2,\nu}\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}},\ \bar{x} - \bar{y} + t_{\alpha/2,\nu}\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}} \right)$$

Independent Samples: *t*-Intervals



Independent Samples: *t*-Intervals

- *Example*: Suppose we have two sets of data

$$\textbf{A}:\ n = 24,\ \ \bar{x} = 9.005,\ s_x = 3.438$$

$$\textbf{B}:\ m = 34,\ \ \bar{y} = 11.864,\ s_y = 3.305$$

What is the 99% two-sided confidence interval?

Independent Samples: *t*-Intervals

- *Example*: Suppose we have two sets of data

$$\textbf{A}:\ n = 24,\ \ \bar{x} = 9.005,\ s_x = 3.438$$

$$\textbf{B}:\ m = 34,\ \ \bar{y} = 11.864,\ s_y = 3.305$$

What is the 99% two-sided confidence interval?

1) First, calculate the degrees of freedom:

$$\nu = \frac{\left(\frac{3.438^2}{24} + \frac{3.305^2}{34}\right)^2}{\frac{3.438^4}{24^2 \times 23} + \frac{3.305^4}{34^2 \times 33}} = 48.43 \quad \rightarrow \mathbf{48}$$

Independent Samples: *t*-Intervals

- *Example*: Suppose we have two sets of data

$$\textbf{A}:\ n = 24,\ \ \bar{x} = 9.005,\ s_x = 3.438$$

$$\textbf{B}:\ m = 34,\ \ \bar{y} = 11.864,\ s_y = 3.305$$

What is the 99% two-sided confidence interval?

1) First, calculate the degrees of freedom:

$$\nu = \frac{\left(\frac{3.438^2}{24} + \frac{3.305^2}{34}\right)^2}{\frac{3.438^4}{24^2 \times 23} + \frac{3.305^4}{34^2 \times 33}} = 48.43 \quad \rightarrow \mathbf{48}$$

2) Second, find the critical point: $t_{0.005,48} = 2.6822$

Independent Samples

- *Example*: Suppose we have two sets of data

$$\mathbf{A}:\ n = 24,\ \ \bar{x} = 9.005,\ s_x = 3.438$$

$$\mathbf{B}:\ m = 34,\ \ \bar{y} = 11.864,\ s_y = 3.305$$

What is the 99% two-sided confidence interval?

3) Then, the interval becomes:

$$\mu_A - \mu_B \in \left(9.005 - 11.864 - 2.6822\sqrt{\frac{3.438^2}{24} + \frac{3.305^2}{34}},\right.$$
$$\left. 9.005 - 11.864 + 2.6822\sqrt{\frac{3.438^2}{24} + \frac{3.305^2}{34}}\right)$$
$$= (-5.28, -0.44)$$

Independent Samples

- *Example*: Suppose we have two sets of data

$$\textbf{A}:\ n = 24,\ \ \bar{x} = 9.005,\ s_x = 3.438$$

$$\textbf{B}:\ m = 34,\ \ \bar{y} = 11.864,\ s_y = 3.305$$

As we can see from the two-sided confidence interval, zero is not included in the interval.

Independent Samples: *t*-Intervals

- **One-sided confidence intervals** are also similar:

$$\mu_A - \mu_B \in \left(-\infty, \bar{x} - \bar{y} + t_{\alpha,\nu} \sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}} \right)$$

and

$$\mu_A - \mu_B \in \left(\bar{x} - \bar{y} - t_{\alpha,\nu} \sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}, \infty \right)$$

- *Recall*: For one sample data

$$\mu \in \left(-\infty, \bar{x} + \frac{t_{\alpha,n-1} s}{\sqrt{n}} \right) \qquad \mu \in \left(\bar{x} - \frac{t_{\alpha,n-1} s}{\sqrt{n}}, \infty \right)$$

Independent Samples: Hypothesis Test

- Consider the **two-sided hypothesis testing** problem:

$$H_0 : \mu_A - \mu_B = \delta \quad \text{versus} \quad H_A : \mu_A - \mu_B \neq \delta$$

- Now the *t*-statistic can be written as

$$t = \frac{\bar{x} - \bar{y} - \delta}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}}$$

- *Recall*: For one-sample data

$$t = \frac{\sqrt{n}(\bar{x} - \mu_0)}{s} = \frac{(\bar{x} - \mu_0)}{s/\sqrt{n}} = \frac{(\bar{x} - \mu_0)}{\mathrm{s.e.}(\hat{\mu})}$$

Independent Samples: Hypothesis Test

- Consider the **two-sided hypothesis testing** problem:

$$H_0 : \mu_A - \mu_B = \delta \quad \text{versus} \quad H_A : \mu_A - \mu_B \neq \delta$$

- When testing if A and B have the same population mean, we set $\delta = 0$

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}}$$

Independent Samples: Hypothesis Test

For the two-sided hypothesis testing problem

$$H_0 : \mu_A - \mu_B = \delta \quad \text{versus} \quad H_A : \mu_A - \mu_B \neq \delta$$

for some fixed value δ of interest (usually $\delta = 0$), the appropriate t-statistic is

$$t = \frac{\bar{x} - \bar{y} - \delta}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}}$$

The two-sided p-value is calculated as

$$\boxed{p\text{-value} = 2 \times P(X > |t|)}$$

where the random variable X has a t-distribution with ν degrees of freedom, and a size α hypothesis test accepts the null hypothesis if

$$\boxed{|t| \leq t_{\alpha/2,\nu}}$$

and rejects the null hypothesis when

$$\boxed{|t| > t_{\alpha/2,\nu}}$$

Independent Samples

- *Example (continued)*: Suppose we have two sets of data

$$\textbf{A}:\ n = 24,\ \ \bar{x} = 9.005,\ s_x = 3.438$$

$$\textbf{B}:\ m = 34,\ \ \bar{y} = 11.864,\ s_y = 3.305$$

How do we do the following hypothesis test?

$$H_0 : \mu_A = \mu_B \quad \text{versus} \quad H_A : \mu_A \neq \mu_B$$

(In other words, the above is the same as saying that $\delta = 0$ in the below test)

$$H_0 : \mu_A - \mu_B = \delta \quad \text{versus} \quad H_A : \mu_A - \mu_B \neq \delta$$

Independent Samples

- *Example (continued)*: Suppose we have two sets of data

$$\textbf{A}:\ n = 24,\ \ \bar{x} = 9.005,\ s_x = 3.438$$

$$\textbf{B}:\ m = 34,\ \ \bar{y} = 11.864,\ s_y = 3.305$$

Hypothesis test:

$$H_0 : \mu_A = \mu_B \quad \text{versus} \quad H_A : \mu_A \neq \mu_B$$

The *t*-statistic is

$$t = \frac{\bar{x} - \bar{y} - \delta}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}} = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}} = \frac{9.005 - 11.864}{\sqrt{\frac{3.438^2}{24} + \frac{3.305^2}{34}}} = -3.169$$

Independent Samples

- *Example (continued)*: For the hypothesis test:

$$H_0 : \mu_A = \mu_B \quad \text{versus} \quad H_A : \mu_A \neq \mu_B$$

The two-sided p-value is therefore

$$p\text{-value} = 2 \times P(X > 3.169)$$

where the random variable X has a t-distribution with degrees of freedom

$$\nu = \frac{\left(\frac{3.438^2}{24} + \frac{3.305^2}{34}\right)^2}{\frac{3.438^4}{24^2 \times 23} + \frac{3.305^4}{34^2 \times 33}} = 48.43$$

Using the integer value $\nu = 48$ gives

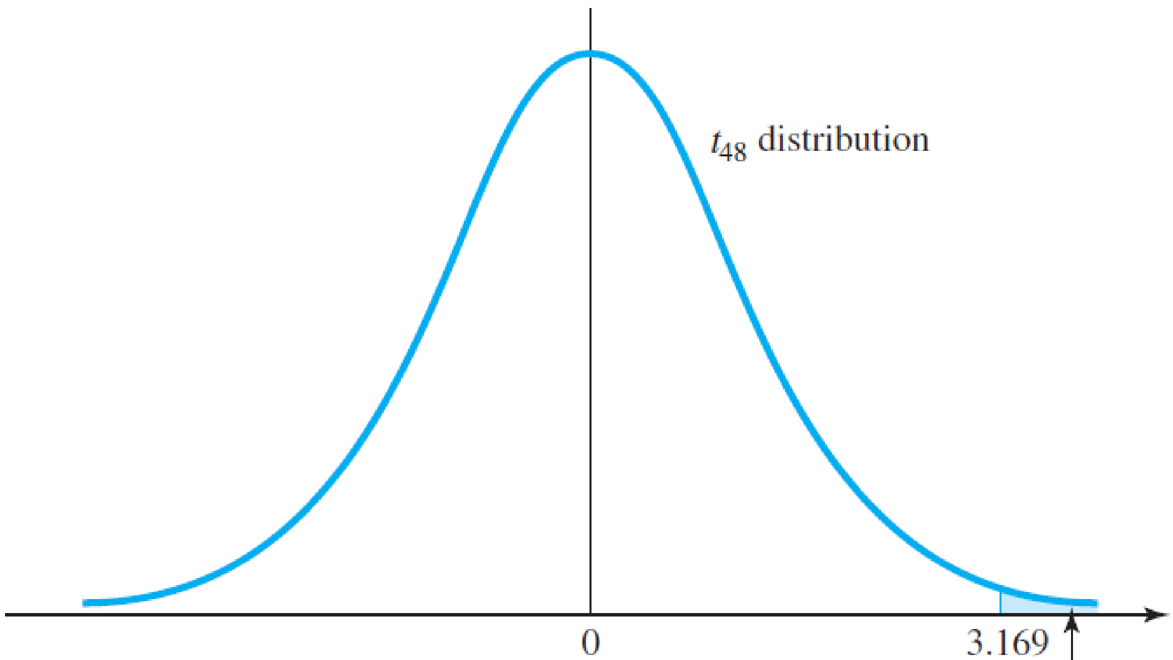
$$p\text{-value} \simeq 2 \times 0.00135 = 0.0027$$

Null hypothesis is rejected and conclude that $\mu_A \neq \mu_B$

Independent Samples

- *Example (continued)*: For the hypothesis test:

$$H_0 : \mu_A = \mu_B \quad \text{versus} \quad H_A : \mu_A \neq \mu_B$$



p-value $= 2 \times P(X > 3.169) = 2 \times 0.00135 = 0.0027$

Independent Samples

- *Example (continued)*: Suppose we have two sets of data

$$\mathbf{A}\text{:}\;\; n = 24,\;\; \bar{x} = 9.005,\; s_x = 3.438$$

$$\mathbf{B}\text{:}\;\; m = 34,\;\; \bar{y} = 11.864,\; s_y = 3.305$$

Recall: The two-sided confidence interval did not include zero.

This is consistent with the hypothesis test rejecting the null hypothesis that the two means are the same.

Independent Samples: Hypothesis Test

A one-sided hypothesis testing problem

$$H_0 : \mu_A - \mu_B \leq \delta \quad \text{versus} \quad H_A : \mu_A - \mu_B > \delta$$

has a p-value

$$p\text{-value} = P(X > t)$$

and a size α hypothesis test accepts the null hypothesis if

$$t \leq t_{\alpha,\nu}$$

and rejects the null hypothesis if

$$t > t_{\alpha,\nu}$$

A size α test for the one-sided hypotheses

$$H_0 : \mu \leq \mu_0 \quad \text{versus} \quad H_A : \mu > \mu_0$$

rejects the null hypothesis when

$$t > t_{\alpha,n-1}$$

and accepts the null hypothesis when

$$t \leq t_{\alpha,n-1}$$

Independent Samples: Hypothesis Test

Similarly, the one-sided hypothesis testing problem

$$H_0 : \mu_A - \mu_B \geq \delta \quad \text{versus} \quad H_A : \mu_A - \mu_B < \delta$$

has a p-value

$$p\text{-value} = P(X < t)$$

and a size α hypothesis test accepts the null hypothesis if

$$t \geq -t_{\alpha,\nu}$$

and rejects the null hypothesis if

$$t < -t_{\alpha,\nu}$$

A size α test for the one-sided hypotheses

$$H_0 : \mu \geq \mu_0 \quad \text{versus} \quad H_A : \mu < \mu_0$$

rejects the null hypothesis when

$$t < -t_{\alpha,n-1}$$

and accepts the null hypothesis when

$$t \geq -t_{\alpha,n-1}$$

Independent Samples

	Hypothesis Testing			Confidence Intervals		
	$H_0 : \mu_A - \mu_B = \delta$ versus $H_A : \mu_A - \mu_B \neq \delta$			$\left(\bar{x} - \bar{y} - t_{\alpha/2,\nu}\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}, \bar{x} - \bar{y} + t_{\alpha/2,\nu}\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}\right)$		
	Significance Levels			Confidence Levels		
p-Value	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$	$1 - \alpha = 0.90$	$1 - \alpha = 0.95$	$1 - \alpha = 0.99$
≥ 0.10	accept H_0	accept H_0	accept H_0	contains δ	contains δ	contains δ
$0.05 - 0.10$	reject H_0	accept H_0	accept H_0	does not contain δ	contains δ	contains δ
$0.01 - 0.05$	reject H_0	reject H_0	accept H_0	does not contain δ	does not contain δ	contains δ
< 0.01	reject H_0	reject H_0	reject H_0	does not contain δ	does not contain δ	does not contain δ

Independent Samples

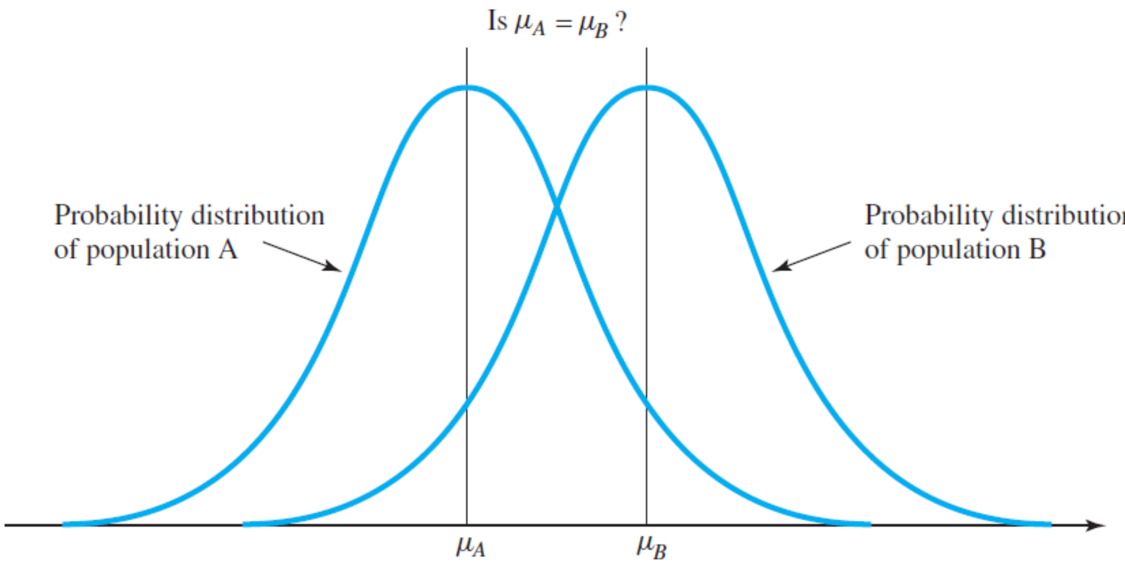
- The procedures with independent samples discussed so far are known as

  two-sample *t*-tests *without* a pooled variance estimate

Independent Samples: Pooled Variance

- What if we want to assume that the two variance are the same?

$$\sigma_A^2 = \sigma_B^2$$



Independent Samples: Pooled Variance

- If we assume that the variances are equal to a common value σ^2, then

$$\sigma_A^2 = \sigma_B^2 = \sigma^2$$

and we can estimate σ^2 with

$$\hat{\sigma}^2 = s_p^2 = \frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}$$

pooled variance estimate

Independent Samples: Pooled Variance

- If we assume that the variances are equal to a common value σ^2, then

$$\sigma_A^2 = \sigma_B^2 = \sigma^2$$

and we can estimate σ^2 with

$$\hat{\sigma}^2 = s_p^2 = \frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}$$

Also, the standard error

$$\text{s.e.}(\bar{x} - \bar{y}) = \sqrt{\frac{\sigma_A^2}{n} + \frac{\sigma_B^2}{m}} = \sigma\sqrt{\frac{1}{n} + \frac{1}{m}}$$

can be estimated as

$$\text{s.e.}(\bar{x} - \bar{y}) = s_p\sqrt{\frac{1}{n} + \frac{1}{m}}$$

Independent Samples: Pooled Variance

- When a pooled variance estimate is employed, p-values and critical points are calculated from a ***t*-distribution with $n + m - 2$ degrees of freedom**

Independent Samples: Pooled Variance

- When a pooled variance estimate is employed, p-values and critical points are calculated from a **t-distribution with $n + m - 2$ degrees of freedom**
- Therefore, a two-sided $1 - \alpha$ level confidence interval for $\mu_A - \mu_B$ is

$$\mu_A - \mu_B \in \left(\bar{x} - \bar{y} - t_{\alpha/2,n+m-2}\, s_p \sqrt{\frac{1}{n} + \frac{1}{m}},\ \bar{x} - \bar{y} + t_{\alpha/2,n+m-2}\, s_p \sqrt{\frac{1}{n} + \frac{1}{m}} \right)$$

$$\mu \in (\hat{\mu} - \text{critical point} \times \text{s.e.}(\hat{\mu}),\ \ \hat{\mu} + \text{critical point} \times \text{s.e.}(\hat{\mu}))$$

Independent Samples: Pooled Variance

- *Example*: Suppose we have two sets of data

$$\textbf{A}:\ n = 24,\ \ \bar{x} = 9.005,\ s_x = 3.438$$

$$\textbf{B}:\ m = 34,\ \ \bar{y} = 11.864,\ s_y = 3.305$$

These are the same data from the previous example

Independent Samples: Pooled Variance

- *Example*: Suppose we have two sets of data

A: $n = 24, \ \bar{x} = 9.005, \ \cancel{s_x = 3.438}$

B: $m = 34, \ \bar{y} = 11.864, \ \cancel{s_y = 3.305}$

$\boxed{s_p = 3.360}$

These are the same data from the previous example

But if we assume the population variances to be the same, the common standard deviation becomes:

$$s_p = \sqrt{\frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}} = \sqrt{\frac{(23 \times 3.438^2) + (33 \times 3.305^2)}{24+34-2}} = 3.360$$

Independent Samples: Pooled Variance

- *Example*: Suppose we have two sets of data

$$\mathbf{A}:\ n = 24,\ \ \bar{x} = 9.005,\ \cancel{s_x = 3.438}$$

$$\mathbf{B}:\ m = 34,\ \ \bar{y} = 11.864,\ \cancel{s_y = 3.305}$$

$$\boxed{s_p = 3.360}$$

How do we do the following hypothesis test?

$$H_0 : \mu_A = \mu_B \quad \text{versus} \quad H_A : \mu_A \neq \mu_B$$

Independent Samples: Pooled Variance

- *Example*: Suppose we have two sets of data

$$\mathbf{A}:\ n = 24,\quad \bar{x} = 9.005,\quad \cancel{s_x = 3.438}$$

$$\mathbf{B}:\ m = 34,\quad \bar{y} = 11.864,\quad \cancel{s_y = 3.305}$$

$$\boxed{s_p = 3.360}$$

How do we do the following hypothesis test?

$$H_0 : \mu_A = \mu_B \quad \text{versus} \quad H_A : \mu_A \neq \mu_B$$

The t-statistic is

$$t = \frac{\bar{x} - \bar{y}}{s_p\sqrt{\frac{1}{n} + \frac{1}{m}}} = \frac{9.005 - 11.864}{3.360\sqrt{\frac{1}{24} + \frac{1}{34}}} = -3.192$$

Independent Samples: Pooled Variance

- *Example*: Suppose we have two sets of data

$$\textbf{A}: \; n = 24, \;\; \bar{x} = 9.005, \;\; \cancel{s_x = 3.438}$$

$$\textbf{B}: \; m = 34, \;\; \bar{y} = 11.864, \;\; \cancel{s_y = 3.305}$$

$\boxed{s_p = 3.360}$

How do we do the following hypothesis test?

$$H_0 : \mu_A = \mu_B \quad \text{versus} \quad H_A : \mu_A \neq \mu_B$$

The t-statistic is

$$t = \frac{\bar{x} - \bar{y}}{s_p \sqrt{\frac{1}{n} + \frac{1}{m}}} = \frac{9.005 - 11.864}{3.360\sqrt{\frac{1}{24} + \frac{1}{34}}} = -3.192$$

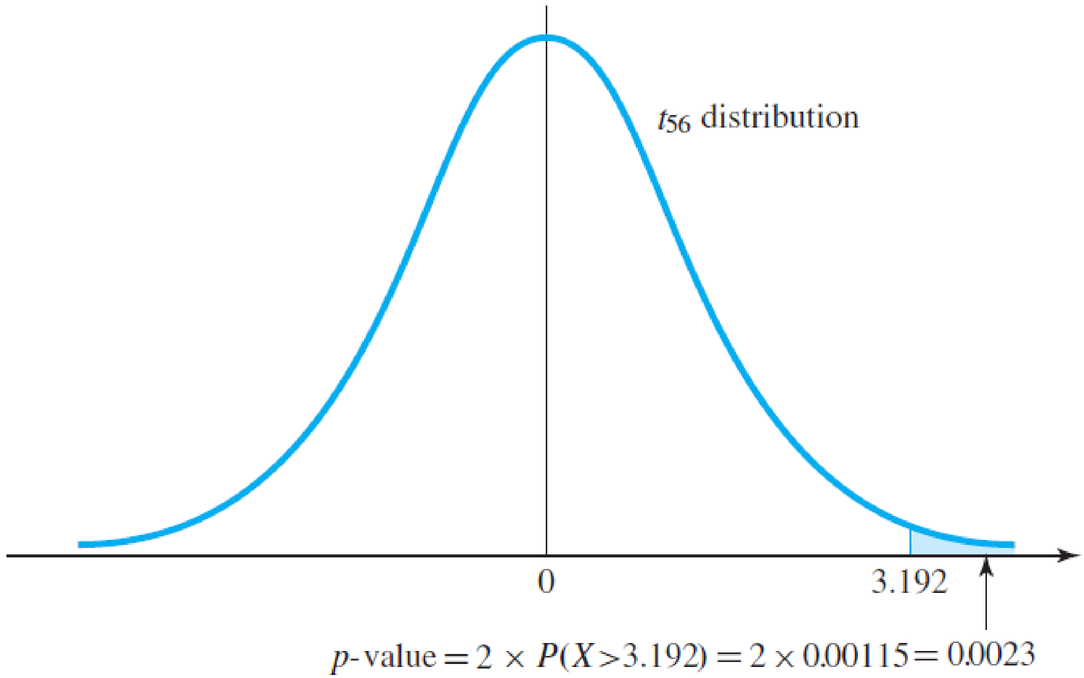
and the p-value is $\quad p\text{-value} = 2 \times P(X > 3.192) \simeq 2 \times 0.00115 = 0.0023$

where the random variable X has a t-distribution with degrees of freedom $n + m - 2 = 56$,

Independent Samples: Pooled Variance

- *Example*: Hypothesis test

$$H_0 : \mu_A = \mu_B \quad \text{versus} \quad H_A : \mu_A \neq \mu_B$$



Independent Samples: Pooled Variance

- *Example*: Suppose we have two sets of data

$$\mathbf{A}:\ n = 24,\ \ \bar{x} = 9.005,\ \cancel{s_x = 3.438}$$

$$\mathbf{B}:\ m = 34,\ \ \bar{y} = 11.864,\ \cancel{s_y = 3.305}$$

$$\boxed{s_p = 3.360}$$

What about a 99% two-sided confidence interval?

Independent Samples: Pooled Variance

- *Example*: Suppose we have two sets of data

$$\mathbf{A}: \; n = 24, \;\; \bar{x} = 9.005, \;\; \cancel{s_x = 3.438}$$

$$\mathbf{B}: \; m = 34, \;\; \bar{y} = 11.864, \;\; \cancel{s_y = 3.305}$$

$$\boxed{s_p = 3.360}$$

What about a 99% two-sided confidence interval?

First, find the critical point: $t_{0.005,56} = 2.6665$

Independent Samples: Pooled Variance

- *Example*: Suppose we have two sets of data

A: $n = 24$, $\bar{x} = 9.005$, $\cancel{s_x = 3.438}$

B: $m = 34$, $\bar{y} = 11.864$, $\cancel{s_y = 3.305}$

$\boxed{s_p = 3.360}$

What about a 99% two-sided confidence interval?

First, find the critical point: $t_{0.005,56} = 2.6665$

Then, the two-sided interval becomes:

$$\mu_A - \mu_B \in \left(9.005 - 11.864 - 2.6665 \times 3.360 \times \sqrt{\frac{1}{24} + \frac{1}{34}},\right.$$

$$\left. 9.005 - 11.864 + 2.6665 \times 3.360 \times \sqrt{\frac{1}{24} + \frac{1}{34}}\right)$$

$$= (-5.25, -0.47)$$

Independent Samples: Pooled Variance

A two-sided p-value is calculated as $2 \times P(X > |t|)$, where the random variable X has a t-distribution with $n + m - 2$ degrees of freedom, and one-sided p-values are $P(X > t)$ and $P(X < t)$. A size α two-sided hypothesis test accepts the null hypothesis if

$$|t| \leq t_{\alpha/2,n+m-2}$$

and rejects the null hypothesis when

$$|t| > t_{\alpha/2,n+m-2}$$

and size α one-sided hypothesis tests have rejection regions $t > t_{\alpha,n+m-2}$ or $t < -t_{\alpha,n+m-2}$.

These procedures are known as two-sample t-tests *with* a pooled variance estimate.

Independent Samples

- Pooled or un-pooled?
  - General procedure is without a pooled variance estimate (this is the safer choice)
  - However, the two population variances are known to be equal, then the pooled variance provides a powerful analysis (e.g., shorter confidence interval)
  - If the results from with and without pooled variance estimates are different, that is likely because the two sample standard deviations are different

Two Sample *z*-Procedures

- If we assume that the population variances σ_A^2 and σ_B^2 are known, then we can use the standard normal distribution instead of a *t*-distribution
- Similar to the one-sample case, two-sample *t*-tests with large sample sizes are essentially equivalent to the two-sample *z*-tests
  - Thus, two-sample *t*-tests can be considered as small-sample tests

Two Sample z-Procedures

A two-sided $1-\alpha$ level confidence interval for the difference in population means $\mu_A - \mu_B$ is

$$\mu_A - \mu_B \in \left(\bar{x} - \bar{y} - z_{\alpha/2}\sqrt{\frac{\sigma_A^2}{n} + \frac{\sigma_B^2}{m}}, \bar{x} - \bar{y} + z_{\alpha/2}\sqrt{\frac{\sigma_A^2}{n} + \frac{\sigma_B^2}{m}} \right)$$

The appropriate z-statistic for the null hypothesis $H_0 : \mu_A - \mu_B = \delta$ is

$$z = \frac{\bar{x} - \bar{y} - \delta}{\sqrt{\frac{\sigma_A^2}{n} + \frac{\sigma_B^2}{m}}}$$

A two-sided p-value is calculated as $2 \times \Phi(-|z|)$, and one-sided p-values are $1 - \Phi(z)$ and $\Phi(z)$. A size α two-sided hypothesis test accepts the null hypothesis if

$$|z| \leq z_{\alpha/2}$$

and rejects the null hypothesis when

$$|z| > z_{\alpha/2}$$

and size α one-sided hypothesis tests have rejection regions $z > z_\alpha$ or $z < -z_\alpha$.

Summary of Chapter 9

- Comparison of two population means
  - Paired and unpaired samples
  - Mainly use *t*-distributions
- Paired samples
  - Take the difference and create a one-sample dataset
  - Then, perform analyses using methods in Chapter 8
- Unpaired samples
  - Unpooled variance: assume that the two datasets have different variances
  - Pooled variance: assume that the two datasets have the same variance
- Use z-procedures if population variances are known

Independent Samples

- More examples …

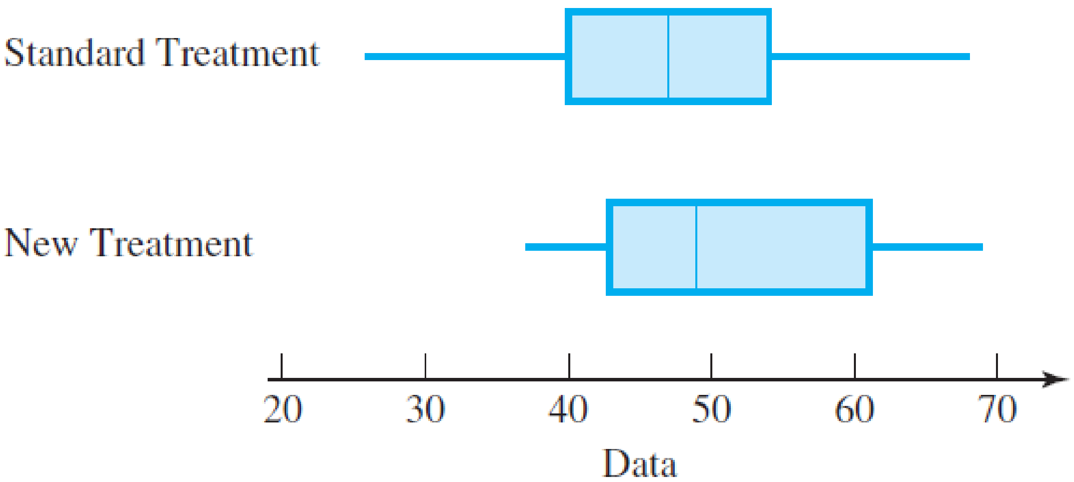
Example: Unpaired Samples

- Data from standard and new treatments are shown below
- Same sample size but unpaired data

Standard treatment x_i	New treatment y_i
33	65
54	61
62	37
46	47
52	45
42	53
34	53
51	69
26	49
68	42
47	40
40	67
46	46
51	43
60	51

Example: Unpaired Samples

- If we simply analyze each data set independently, it is difficult to know whether the two are different with statistical significance



Standard Treatment
Sample size = 15
Sample mean = 47.47
Sample standard deviation = 11.40

New Treatment
Sample size = 15
Sample mean = 51.20
Sample standard deviation = 10.09

Example: Unpaired Samples

- First, let's perform the general (unpooled) procedure
- We want evidence that the new treatment is better than the standard treatment, thus we set the hypothesis test as:

$$H_0 : \mu_A \geq \mu_B \quad \text{versus} \quad H_A : \mu_A < \mu_B$$

Example: Unpaired Samples

- First, let's perform the general (unpooled) procedure
- We want evidence that the new treatment is better than the standard treatment, thus we set the hypothesis test as:

$$H_0 : \mu_A \geq \mu_B \quad \text{versus} \quad H_A : \mu_A < \mu_B$$

- Degrees of freedom:

$$\nu = \frac{\left(\frac{11.40^2}{15} + \frac{10.09^2}{15}\right)^2}{\frac{11.40^4}{15^2 \times 14} + \frac{10.09^4}{15^2 \times 14}} = 27.59$$

which can be rounded down to $\nu = 27$

Example: Unpaired Samples

- First, let's perform the general (unpooled) procedure
- We want evidence that the new treatment is better than the standard treatment, thus we set the hypothesis test as:

$$H_0 : \mu_A \geq \mu_B \quad \text{versus} \quad H_A : \mu_A < \mu_B$$

- Degrees of freedom:

$$\nu = \frac{\left(\frac{11.40^2}{15} + \frac{10.09^2}{15}\right)^2}{\frac{11.40^4}{15^2 \times 14} + \frac{10.09^4}{15^2 \times 14}} = 27.59$$

which can be rounded down to $\nu = 27$

- t-statistic:

$$t = \frac{47.47 - 51.20}{\sqrt{\frac{11.40^2}{15} + \frac{10.09^2}{15}}} = -0.949$$

Example: Unpaired Samples

- Finally, the p-value is:

$$p\text{-value} = P(X < -0.949) = 0.175$$

where the random variable X has a t-distribution with $\nu = 27$ degrees of freedom

➔ Cannot reject the null hypothesis

Example: Unpaired Samples

- For a 99% confidence interval, we first find the critical point

$$t_{0.01,27} = 2.473$$

- Then the interval becomes

$$\mu_A - \mu_B \in \left(-\infty, 47.47 - 51.20 + 2.473\sqrt{\frac{11.40^2}{15} + \frac{10.09^2}{15}}\right)$$
$$= (-\infty, 5.99)$$

➔ The standard treatment may be 6 points better than the new treatment on average

Example: Unpaired Samples

- Suppose we decide to perform a pooled variance analysis

(assume $s_x = 11.40$ and $s_y = 10.09$ are similar)

Example: Unpaired Samples

- Suppose we decide to perform a pooled variance analysis
- For the hypothesis test:

$$H_0 : \mu_A \geq \mu_B \quad \text{versus} \quad H_A : \mu_A < \mu_B$$

- Pooled variance is:

$$s_p^2 = \frac{(14 \times 11.40^2) + (14 \times 10.09^2)}{28} = 115.88$$

so that the pooled standard deviation is $s_p = \sqrt{115.88} = 10.76$

Example: Unpaired Samples

- Suppose we decide to perform a pooled variance analysis
- For the hypothesis test:

$$H_0 : \mu_A \geq \mu_B \quad \text{versus} \quad H_A : \mu_A < \mu_B$$

- Pooled variance is:

$$s_p^2 = \frac{(14 \times 11.40^2) + (14 \times 10.09^2)}{28} = 115.88$$

so that the pooled standard deviation is $s_p = \sqrt{115.88} = 10.76$

- t-statistic becomes:

$$t = \frac{47.47 - 51.20}{10.76\sqrt{\frac{1}{15} + \frac{1}{15}}} = -0.946$$

Example: Unpaired Samples

- Then the p-value for the pooled variance procedure is:

$$p\text{-value} = P(X < -0.946) = 0.175$$

where the random variable X has a t-distribution with $\nu = 28$ degrees of freedom

- Both the unpooled and pooled analyses results in similar p-values
- One-sided 99% confidence interval also turns out to be very similar

Excel Examples

- Next, let's show how to solve similar questions in Excel