Lecture 8
Comparing Population Means

Paired data

Independent data

In the Last Lecture….

- We worked through hypothesis tests for a population mean to answer research questions about a target population.
- We tested a **null hypothesis**: $\mathbf{H_0: \mu = \mu_0}$ against an **alternative hypothesis**: $\mathbf{H_1: \mu \neq \mu_0}$
- We used a z-test when σ was known: $z = \frac{\bar{y}-\mu}{\sigma/\sqrt{n}}$

  We used a t-test with $n-1$ df when σ was unknown: $t = \frac{\bar{y}-\mu}{s/\sqrt{n}}$
- We checked test **assumptions** to ensure the **test-statistic** was valid:

  z-test: check sample mean is from a normal distribution (when sample size is large, CLT for sample means applies)

  t-test: check sample is drawn from a normal distribution (when sample size is large, t-distributions will be robust for non-normality)
- We used a **p-value** (area in the two tails of the distribution of the test statistic) to make decisions.
- Based on these **decisions** we wrote **conclusions** to answer the research question about the target population.

Steps in Hypothesis Testing

Hypothesis Test

H

Hypotheses:

State the *null* and the *alternative* hypotheses in terms of the *parameter* of interest.

A

Assumptions:

Check the underlying assumptions of the test.

T

Test Statistic:

Calculate the test statistic.

P

p-value:

Obtain the p-value for the test from the distribution of the test statistic.

D

Decision:

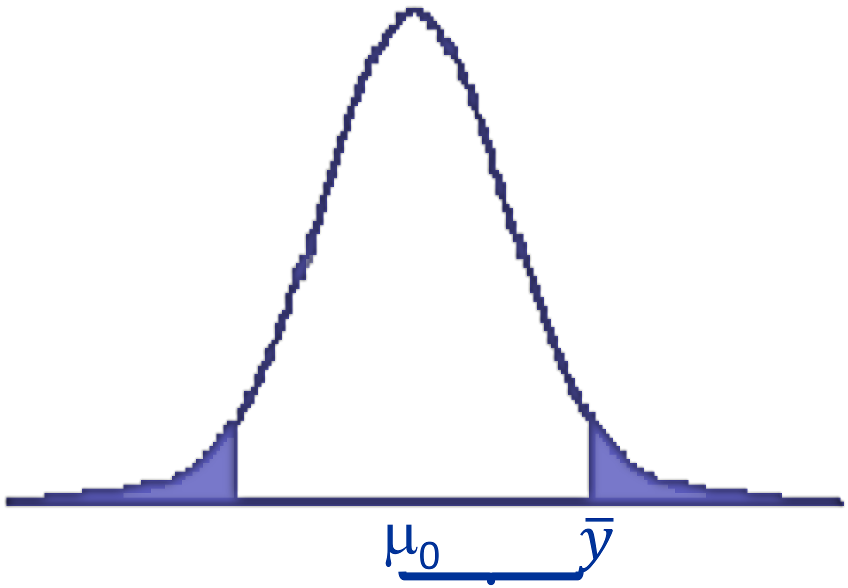
If the p-value is less than 0.05 (the significance level), reject the null hypothesis. If the p-value is not less than 0.05, do not reject the null hypothesis.

C

Conclusion:

Write a conclusion to the original research question in terms of the target population.

When the p-value is Small



The discrepancy between $\bar{y}$ and μ_0 is large

⇓

z is large

⇓

p-value is small

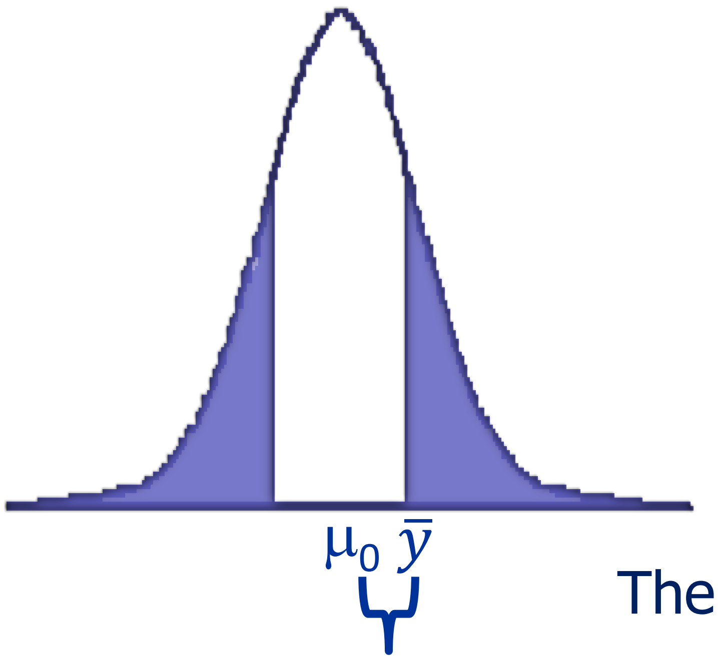
⇓

There is evidence against H_0

⇓

Reject H_0

When the p-value is Large



The discrepancy between $\bar{y}$ and μ_0 is small

z is small

p-value is large

There is no evidence against H_0

Do not reject H_0

Review Quiz 1

After performing a hypothesis test:

$$H_0: \mu = 100$$
$$H_1: \mu \neq 100$$

H_0 is not rejected. State whether the following statements are *True* or *False:*

a. you must have a large test statistic

b. you have a statistically significant result

c. you must have a large p-value

d. the population mean could be equal to 100, the discrepancy between your sample mean and 100 could be due to sampling error

Review Quiz 2

For the following research questions, write down appropriate null and the alternative hypotheses, and the test statistic you would use in each case.

a. Is the average speed at which motorists drive over the Sydney Harbour Bridge 60km per hour, assuming speeds on the bridge follow a normal distribution?

b. Is the average age at which babies are immunised against measles 16 months? It is known that immunisation times follow a normal distribution with a standard deviation of 3 months.

Paired t-test

Beginners' Bootcamp

We will start off with a familiar sort of problem:

Beginners' Bootcamp is a fitness training programme. Participants meet at the local park at 5am each morning for four weeks to work out. A random sample of 10 people are selected after the first week and their weight loss (or gain) is recorded below:

Weight Loss for Beginners' Bootcamp										
Participant	Alex	Bob	Calvin	Diana	Eric	Felix	Gaby	Henry	Ian	Jack
Difference (kg)	4	1	2	0	-1	2	1	3	0	1

On average, was there any significant change in weight after one week on the programme?

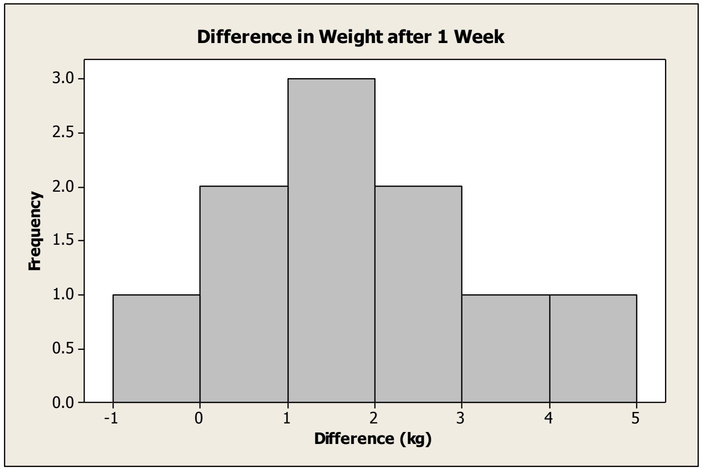
Some Information about the Sample

Some Sample Information:

Variable of interest:

Y = Difference in weight after one week

$n = 10 \quad \bar{y} = 1.3kg \quad s = 1.494kg$



Name	Difference
Alex	4
Bob	1
Calvin	2
Diana	0
Eric	-1
Felix	2
Gaby	1
Henry	3
Ian	0
Jack	1

One Sample t-test for a Population Mean

Hypothesis Test

H $H_0: \mu = 0$

$H_1: \mu \neq 0$

A From the histogram, the differences appear to be drawn from a normal distribution

T $t = \frac{\bar{y}-\mu}{s/\sqrt{n}} = \frac{1.3-0}{1.494/\sqrt{10}} = 2.75$ with $n - 1 = 9\ df$

P $0.02 < \text{p-value} < 0.05$

D Since p-value < 0.05, reject H_0

C There is evidence to suggest that on average, there is a significant difference in weight after one week. It appears that participants lose a significant amount of weight in the first week.

The Differences Between Matched Pairs

Calculating the differences:

We have one sample of participants. For each participant in the sample, we have two measurements – weight before the programme and weight after one week on the programme. We have matched/paired up the before and after weights for each participant. **Each 'before weight' can be matched with a specific 'after weight' by pairing up the two sets of measurements on the participants. This one sample t-test on the differences between matched pairs is otherwise known as a 'paired t-test'.**

Name	Weight Before (kg)	Weight After (kg)	Difference
Alex	94	90	4
Bob	86	85	1
Calvin	77	75	2
Diana	60	60	0
Eric	89	90	-1
Felix	83	81	2
Gaby	59	58	1
Henry	78	75	3
Ian	88	88	0
Jack	93	92	1

Paired Data

- Many research questions take the following form:

  "Does the population mean change after subjects have been given some treatment?"

  eg. "***Do pulse rates change after exercise?"*** We could take *one sample of students and record each student's pulse before and after exercise.*

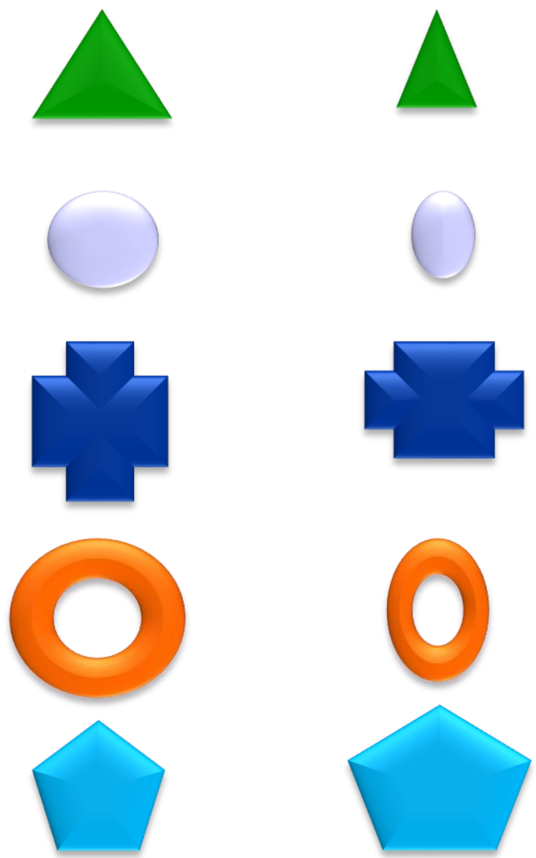
  or ***"Is there a difference between smash repair quotes at Garage A and Garage B?"*** We could take *one sample of cars* and obtain a quote for each car at Garage A and at Garage B.

- Other questions may involve comparing responses for matched units. eg. "***Do mothers and fathers spend a similar amount of time helping children with homework?"*** We could take *one sample of school children* and ask each child's mother and father how many hours per week they spend helping the child with homework.

- ***In each case we have one sample with two matched/paired measurements on each subject. We match up pairs for each subject and use a one sample test on the differences between pairs.***

Differences Between Matched Pairs

Subject	Obs1	Obs2	Difference
A			
B			
C			
D			
E			



We carry out a (one sample) test on the differences to see if the average difference could be 0:

$$H_0:\ \mu_d = 0$$

Paired t-test

Null and alternative hypotheses for paired t-test:

$H_{0:}\ \mu_d = 0,\ \ H_{1:}\ \mu_d \neq 0$

Test statistic for paired t-test:

with n – 1 degrees of freedom

(where n is the number of differences)

$$t = \frac{\bar{y}_d - \mu_d}{s_d / \sqrt{n_d}}$$

where $\bar{y}_d$ is the mean of the differences (calculate from the sample of differences) and s_d is the standard deviation of the sample of differences.

Example of a Paired t-test: Some Sample Information

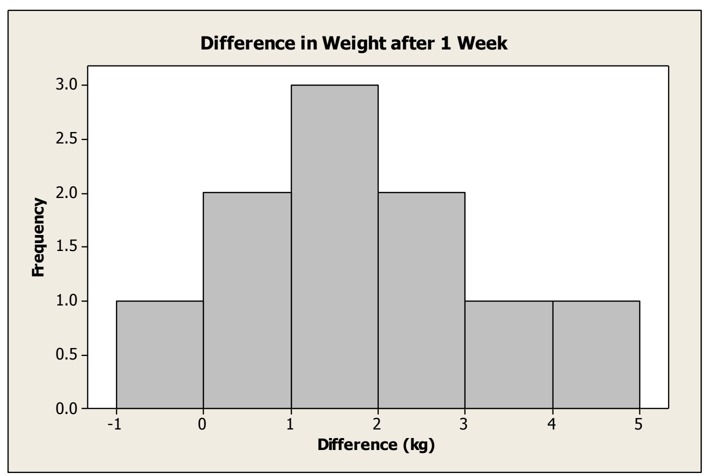
Some Sample Information:

BOOT CAMP

Variable of interest:

Y_d = Difference in weight after one week

$n_d = 10$ $\quad \bar{y}_d = 1.3\text{kg}$ $\quad s_d = 1.494\text{kg}$



Name	Difference
Alex	4
Bob	1
Calvin	2
Diana	0
Eric	-1
Felix	2
Gaby	1
Henry	3
Ian	0
Jack	1

Paired t-test: Bootcamp

Hypothesis Test

H H_0: $\mu_d = 0$

H_1: $\mu_d \neq 0$

A From the histogram, the differences appear to be drawn from a normal distribution

T $t = \frac{\bar{y}_d - \mu_d}{s_d / \sqrt{n_d}} = \frac{1.3 - 0}{1.494 / \sqrt{10}}$ = 2.75 with $n_d - 1 = 9\ df$

P $0.02 < \text{p-value} < 0.05$

D Since p-value < 0.05, reject H_0

C There is evidence to suggest that on average, there is a significant difference in weight after one week. It appears that participants lose a significant amount of weight in the first week.

95% Confidence Interval for μ_d

To calculate a 95% confidence interval to estimate the average difference between matched pairs in the target population:

$95\%\ CI\ for\ \mu_d = \bar{y} \pm t_{crit} \times {}^{s_d}/_{\sqrt{n}}$ where t_{crit} is the critical value which cuts off 5% in the two tails of the t-distribution with $n-1\ df$

To estimate the difference in weight before and after one week in Bootcamp for the target population:

$$95\%\ CI\ for\ \mu_d = 1.3 \pm 2.262 \times {}^{1.494}/_{\sqrt{10}} = (0.231, 2.369)$$

We can be 95% confident that the average weight of participants before Bootcamp is between 0.23kg and 2.37kg heavier than the average weight after one week on Bootcamp.

Minitab Output

We will get this Minitab output if we enter the differences into a single column and use the one sample t-test option

```
One-Sample T: Difference
Test of mu = 0 vs not = 0

Variable     N   Mean  StDev  SE Mean       95% CI            T      P
Difference  10  1.300  1.494    0.473  (0.231, 2.369)  2.75  0.022
```

Minitab gives exact p-values for t-tests

```
Paired T-Test and CI: Weight Before, Weight After
Paired T for Weight Before - Weight After

                N   Mean  StDev  SE Mean
Weight Before  10  80.70  12.49     3.95
Weight After   10  79.40  12.31     3.89
Difference     10  1.300  1.494    0.473

95% CI for mean difference: (0.231, 2.369)
T-Test of mean difference = 0 (vs not = 0): T-Value = 2.75  P-Value = 0.022
```

We will get this Minitab output if we enter the measurements into two columns - before weights in one column and after weights in another then use the paired t-test option

Paired t-test

Hypothesis Test

H

H_0: $\mu_d = 0$ (Differences are from a population with mean = 0, ie. No difference between population means)

H_1: $\mu_d \neq 0$ (Population means are not equal)

A

From the histogram, the differences appear to be drawn from a normal distribution………

T

P

D

C

So a paired t-test is just a one sample test for a population mean where the variable of interest consists of the differences between matched pairs.

Quiz 4

Research Question: In 2011, were grocery prices in Melbourne significantly different from grocery prices in Brisbane?

A group of Introductory Statistics students in 2011 conducted a study to compare grocery prices in Melbourne and Brisbane. The students selected a sample of 13 grocery items to compare in the two cities. They visited a number of grocery stores in each city to ascertain the typical retail price of each item in each city. The data on the following slide show the typical price of each item in both Melbourne and Brisbane. Carry out an appropriate hypothesis test, in the space provided, to answer the research question.

Source: Black et al, *Australasian Business Statistics*, (2013), Wiley

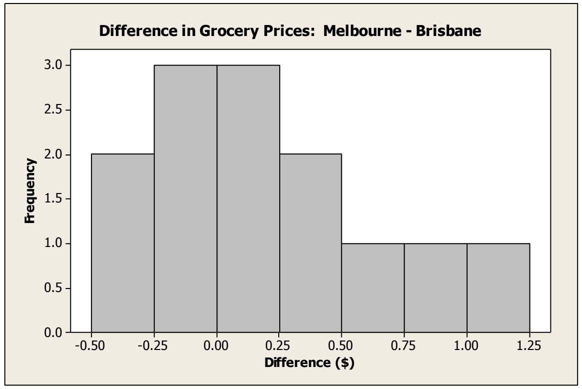
Quiz 4

Research Question: In 2011, were grocery prices in Melbourne significantly different from grocery prices in Brisbane?

Item	Melbourne	Brisbane	
Milk, 2 litre whole milk	$2.85	$2.55	
Bread, while sliced (650-750g)	$3.50	$3.15	
Roast beef (1 kg)	$10.70	$11.16	
Bacon, middle rashers (1 kg)	$10.29	$9.77	
Bananas (1 kg)	$12.87	$11.98	
Potatoes (1 kg)	$2.60	$2.42	
Carrots (1 kg)	$2.11	$2.37	
Pineapple, sliced (450 g can)	$2.12	$2.06	
Chocolate, milk (200 g block)	$4.21	$4.31	
Eggs, free range (1 dozen)	$5.29	$5.39	
Jam, strawberry (500 g jar)	$3.30	$3.32	
Laundry detergent (875 g)	$9.53	$8.39	
Toilet tissue (8×180 sheet rolls)	$6.39	$6.29	

Solution to Quiz 4

Item	Melbourne	Brisbane	Difference M-B
Milk,	$2.85	$2.55	0.3
Bread	$3.50	$3.15	0.35
Roast beef	$10.70	$11.16	-0.46
Bacon	$10.29	$9.77	0.52
Bananas	$12.87	$11.98	0.89
Potatoes	$2.60	$2.42	0.18
Carrots	$2.11	$2.37	-0.26
Pineapple	$2.12	$2.06	0.06
Chocolate	$4.21	$4.31	-0.10
Eggs	$5.29	$5.39	-0.10
Jam	$3.30	$3.32	-0.02
Laundry detergent	$9.53	$8.39	1.14
Toilet tissue	$6.39	$6.29	0.1



Solution to Quiz 4

Quiz 5

Calculate and interpret a 95% confidence interval to estimate the difference between grocery prices in Melbourne and Brisbane:

Two sample t-test

Two Sample t-test for Comparing Population Means

- ***Now we will consider two sample problems.*** Here we consider research questions which are concerned with the difference between populations means for two independent populations. For example:
  - *"Was the average age of passengers on cruise ships different in 2004 to the average age of passengers in 2014?*
  - *"Do customers who pay restaurant bills using a credit card tip any more or less than customers who pay with cash?"*
- We could not use paired t-tests to address either of these questions as there is no meaningful way to match up subjects in the two samples. Each question compares the average score in two independent populations. These questions could be answered using two-sample t-tests.

Two Sample t-test for Comparing Population Means

Research Question: Does the type of tag used to identify penguins affect the time spent foraging for food?

Scientists trying to tell penguins apart have several ways to tag the birds. One method involves wrapping metal strips with ID numbers around the penguin's flipper, while another involves electronic tags. Scientists are concerned that the method used to tag the penguins may affect the time they spend foraging for food. Longer foraging trips can jeopardise both breeding success and survival of chicks waiting for food. Data were collected from a sample of penguins that were randomly given either metal or electronic tags. From this study we have information on foraging times, in days, spent on 30 foraging trips for the penguins with metal tags and another 28 foraging trips for the penguins with electronic tags. We will use a two sample t-test on these data to address the research question above.

Source: adapted from Lock et al, *Statistics Unlocking the Power of Data* (2013), Wiley

Metal Tags vs Electronic Tags

Metal Tags – Days		Electronic Tags – Days	
5.2	13.8	3.9	11.1
5.3	14.0	5.6	11.8
7.5	14.1	5.7	12.1
9.2	14.5	6.0	13.4
9.3	15.1	6.7	14.2
9.8	15.8	6.7	14.3
10.1	16.0	7.4	14.6
10.7	16.1	7.4	14.8
10.8	16.3	8.1	15.7
10.8	16.5	8.6	16.2
11.7	16.6	9.2	19.5
11.8	18.4	9.7	19.6
13.0	19.3	10.1	20.9
13.6	19.7	10.2	
13.8	20.0	10.4	

Descriptive Statistics: Metal Tags, Electronic Tags

Variable	N	Mean	StDev	Min	Median	Max
Metal Tags	30	13.30	3.92	5.2	13.8	20.0
Electronic Tags	28	11.21	4.58	3.9	10.3	20.9

A Two-Sample t-test

- ***We have two samples here.*** The samples are drawn from two independent populations. We have taken one measurement from each subject in each sample.
- We want to determine whether the average duration for foraging trips differs for penguins with metal tags and penguins with electronic tags.
- The **null hypothesis** will be that there is **no difference** between the two population mean times.
- The **alternative hypothesis** is that there is **a difference** between the two population means.
- Since we have **two independent samples** from **two independent populations**. There is no meaningful way of pairing up the two samples so we will perform a **two sample t-test.**

Some Information about the Samples

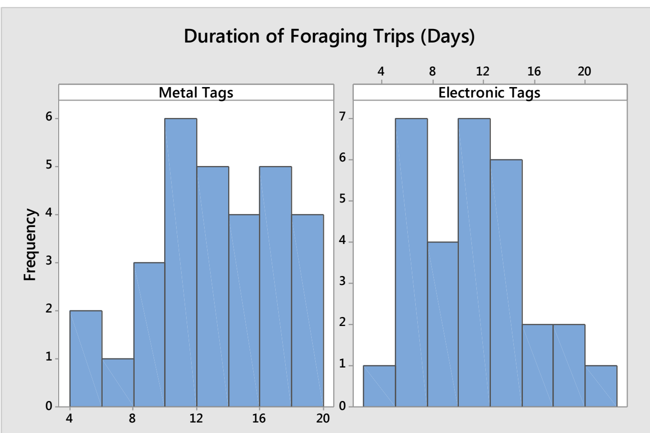
Some Information about the Samples:

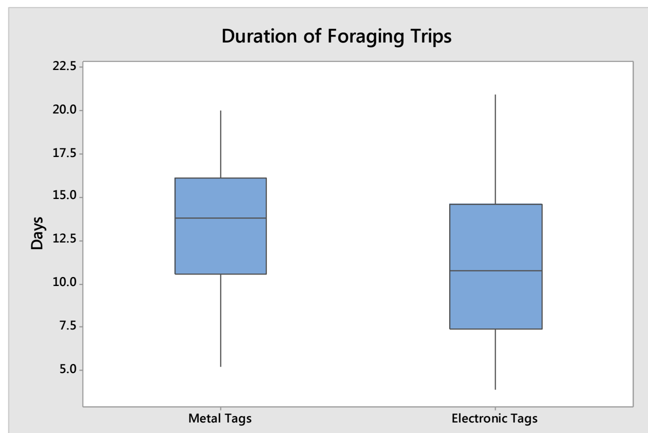
Variables:

Y_M = number of days spent foraging for penguins with metal tags

Y_E = number of days spent foraging for penguins with electronic tags

Or Variables: Y = duration of trip in days, X = type of tag





Descriptive Statistics: Metal Tags, Electronic Tags

Variable	N	Mean	StDev	Min	Median	Max
Metal Tags	30	13.30	3.92	5.2	13.8	20.0
Electronic Tags	28	11.21	4.58	3.9	10.3	20.9

Quiz 6

Research Question:
Is there a difference between the average time spent foraging for penguins with metal tags and penguins with electronic tags?

Before we work through the two sample t-test write a paragraph addressing the research question just using the information about the two samples given on the previous slide:

A Two Sample t-test

Hypothesis Test

H

H_0: $\mu_1 = \mu_2$

H_1: $\mu_1 \neq \mu_2$

Our test is only valid if we can reasonably assume that both samples are drawn from normal distributions

A

AND

We can assume that both samples are drawn from populations with equal spread (ie. we assume $\sigma_1 = \sigma_2$)

T

$t = \frac{\bar{y}_1 - \bar{y}_2}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ with df $= n_1 + n_2 - 2$

Calculating s_p, the Pooled Standard Deviation

- The pooled standard deviation is a weighted average of the two sample standard deviations
- We use a pooled standard deviation because we are assuming the two samples are from populations with the same standard deviation ie. we assume $\sigma_1 = \sigma_{2.}$
- So we calculate s_p to estimate σ_P where σ_P is the common population standard deviation.
- $s_p = \sqrt{\frac{s_1^2(n_1-1)+s_2^2(n_2-1)}{n_1+n_2-2}}$

Quiz 7

Descriptive Statistics: Metal Tags, Electronic Tags

Variable	N	Mean	StDev	Min	Median	Max
Metal Tags	30	13.30	3.92	5.2	13.8	20.0
Electronic Tags	28	11.21	4.58	3.9	10.3	20.9

We noted that the boxplots for the penguins example indicated that the two samples may be drawn from populations with a common population standard deviation, σ_p.

Calculate $s_p = \sqrt{\frac{s_1^2(n_1-1)+s_2^2(n_2-1)}{n_1+n_2-2}}$, to estimate σ_p:

$s_p =$

Two Sample t-test – Metal Tags vs Electronic Tags

Hypothesis Test

H $H_0: \mu_M = \mu_E$ $H_1: \mu_M \neq \mu_E$

A Histograms suggested that both samples were drawn from normal distributions and boxplots indicated that the spread in each population could be equal, as required.

T
$$t = \frac{\bar{y}_M - \bar{y}_E}{s_p\sqrt{\frac{1}{n_M} + \frac{1}{n_E}}} = \frac{13.30 - 11.21}{4.25\sqrt{\frac{1}{30} + \frac{1}{28}}} = 1.87$$

with df $= n_M + n_E - 2 = 30 + 28 - 2 = 56$

P $0.05 < \text{p-value} < 0.1$

D Since p-value ≥ 0.05, do not reject H_0

C There is no evidence to suggest that average duration of foraging trips differs for penguins with metal tags and penguins with electronic tags. The mean times could be the same.

95% Confidence Interval for μ_1-μ_2

To calculate a 95% confidence interval to estimate the average difference between two populations:

$$95\%\ CI\ for\ \mu_1 - \mu_2 = (\bar{y}_1 - \bar{y}_2) \pm t_{crit} \times s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

where t_{crit} is the critical value which cuts off 5% in the two tails of the t-distribution with $n_1 + n_2 - 2\ df$

To estimate the difference in the duration of foraging trips for penguins with metal tags and penguins with electronic tags:

$$95\%\ CI\ for\ \mu_M - \mu_E = (13.30 - 11.21) \pm 2.009 \times 4.25 \sqrt{\frac{1}{30} + \frac{1}{28}} = (-0.15, 4.32)$$

We can be 95% confident that the mean duration of foraging trips for penguins with a metal tags is between 0.15 days less and 4.32 days more than the mean duration of foraging trips for penguins with electronic tags.

Minitab Output

Two-Sample T-Test and CI: Metal Tags, Electronic Tags

Two-sample T for Metal Tags vs Electronic Tags

	N	Mean	StDev	SE Mean
Metal Tags	30	13.30	3.92	0.72
Electronic Tags	28	11.21	4.58	0.87

Difference = μ (Metal Tags) - μ (Electronic Tags)
Estimate for difference: 2.08
95% CI for difference: (-0.15, 4.32)
T-Test of difference = 0 (vs ≠): T-Value = 1.87 P-Value = 0.067 DF = 56
Both use Pooled StDev = 4.2492

Quiz 8

Research Question:
Does the mean length of cuckoos' eggs vary depending on the species of foster parents?

Cuckoos lay their eggs in the nests of other host birds. The eggs are then adopted and hatched by the host birds.

The researchers investigated a number of host birds to determine whether the sizes of cuckoo eggs vary, depending on the species of the foster parents. Eggs were randomly selected from the nests of host birds.

Source: Tippett, L.H.C. (1952), *The Methods of Statistics,* John Wiley & Sons
Drawn from the work of O.M. Latter in 1902 (adapted)

Quiz 8

Sparrow	Robin
22.05	21.05
22.85	21.85
23.05	22.05
23.05	22.05
23.05	22.05
23.05	22.25
23.45	22.45
23.85	22.45
23.85	22.65
23.85	23.05
24.05	23.05
25.05	23.05
	23.05
	23.05
	23.25
	23.85

The data consist of the lengths of cuckoos' eggs (in mm) which were laid in the nests of sparrows and robins.

Use an appropriate test to determine whether the mean length of eggs is dependent on the species of the host bird.

Solution to Quiz 8

Sparrow	Robin
22.05	21.05
22.85	21.85
23.05	22.05
23.05	22.05
23.05	22.05
23.05	22.25
23.45	22.45
23.85	22.45
23.85	22.65
23.85	23.05
24.05	23.05
25.05	23.05
	23.05
	23.05
	23.25
	23.85

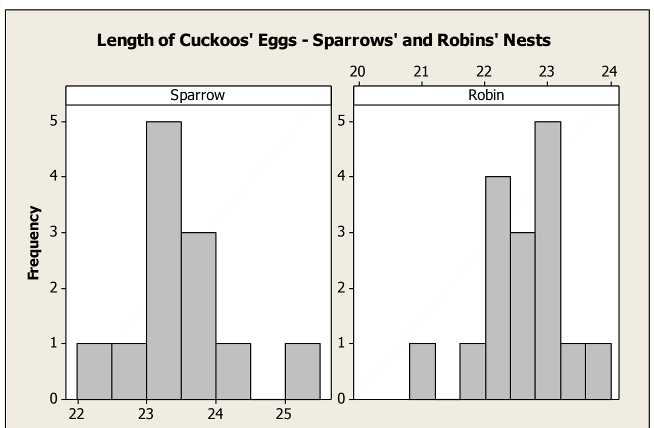
Variables of interest:

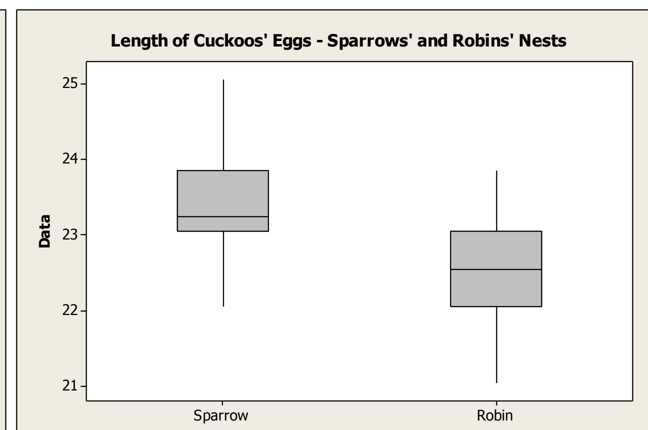
Y_S = length of cuckoos' eggs laid in sparrows' nests

Y_R = length of cuckoos' eggs laid in robins' nests

Descriptive Statistics: Sparrow, Robin

Variable	N	Mean	SE Mean	StDev	Median
Sparrow	12	23.433	0.219	0.760	23.250
Robin	16	22.575	0.171	0.685	22.550





Solution to Quiz 8

Quiz 9

Calculate and interpret a 95% confidence interval to estimate the difference between the average length of cuckoos' eggs laid in sparrows' and robins' nests:

Homework Questions

Don’t forget to try these!!!!!!

Homework Question 1

For the following five problems, determine whether the correct analysis would be a ***paired t-test*** (assuming differences were drawn from a normal distribution) or a ***two sample t-test*** (assuming both samples were drawn from populations which were normally distributed with equal spread).

Homework Question 1a

A random sample of 50 people aged between 25 and 40 years were surveyed and it was found that their average credit debt was $5630 with a standard deviation of $1450. Five years later, another random sample of 50 people in the same age group were surveyed and it was found that their average credit debt was $9340 with a standard deviation of $1870. A hypothesis test was used to determine whether the average credit card debt for people aged 25 to 40 years had changed over the five year period.

Homework Question 1b

In a study to determine whether gender is related to salary offers for graduating MBA students, 25 pairs of students were selected. Each pair consisted of a female and a male student who were matched according to their GPA, units studied, ages and previous work experience. The highest salary offered to each graduate was recorded and a hypothesis test was used to determine whether there was sufficient evidence to infer that gender is a factor in salary offers.

Source: Keller, *Statistics for Management and Economics* (2012), Cengage

Homework Question 1c

Charles Darwin performed an experiment to determine whether self-fertilised and cross-fertilised plants have different growth rates. Pairs of *Zea mays* plants were planted in pots with each pot containing one self-fertilised and one cross-fertilised plant. The heights of the plants were measured after a specified period of time and compared.

Plant height (in 1/8 inches)

Pot	Cross	Self
1	188	139
2	96	163
3	168	160
4	176	160
5	153	147
6	172	149
7	177	149
8	163	122
9	146	132
10	173	144
11	186	130
12	168	144
13	177	102
14	184	124
15	96	144

Source: C Darwin, *The Effects of Cross- and Self-Fertilisation in the Vegetable Kingdom*, D Appleton and Co., New York, 1902.

Homework Question 1d

In a study of interspousal aggression and its possible effect on child behaviour, the behaviour problem checklist (BPC) scores were recorded for 47 children whose parents were classified as aggressive. The sample mean and standard deviation were 7.92 and 3.45 respectively. For a sample of 38 children whose parents were classified as nonaggressive, the mean and standard deviation of the BPC scores were 5.80 and 2.87 respectively. Do these observations indicate that children from aggressive families may differ in their BPC scores to those of children from nonaggressive families?

Source: Johnson, A and Bhattacharyya, G. K., *Statistics, Principles & Methods* (2010), Wiley

Homework Question 1e

The data given below are the winning times (seconds) for the 400 metre freestyle swimming events in the Commonwealth Games for men and women since 1934. Do males typically have different swim times in this event than females?

Year	Men	Women	Year	Men	Women
34	303.00	345.60	78	234.43	248.85
38	294.60	339.70	82	233.29	248.82
50	289.40	326.40	86	232.25	247.68
54	279.80	311.40	90	229.91	248.89
58	265.90	289.40	94	225.77	252.56
62	260.00	291.40	98	224.35	252.39
66	255.00	278.80	02	220.08	249.49
70	248.48	267.38	06	228.17	247.69
74	241.44	262.09	10	228.48	245.68
			14	223.46	244.47

Lecture 8 Summary

- When comparing two sets of measurements to determine whether there is a difference in their population means, we need to ascertain whether the measurements are independent or paired.
- If the measurements are paired we calculate the differences between matched pairs and, assuming differences are from a normal distribution, we do a one-sample t-test on the ONE sample of differences.
- If the measurements are independent then, assuming both samples are from normal distributions with equal spread we use a two-sample t-test where $t = \frac{\bar{y}_1 - \bar{y}_2}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ $with\ df = n_1 + n_2 - 2$

and $s_p = \sqrt{\frac{s_1^2(n_1-1)+s_2^2(n_2-1)}{n_1+n_2-2}}$

Textbook References

Further information on the topics discussed in this lecture can be found in:

Modern Statistics: An Introduction
by Don McNeil and Jenny Middledorp
(ISBN 9781486007011).

- Chapter 8: Pages 162 – 165
  and 167-170

More of this chapter will be covered later in the course.