MATU9D2 : PRACTICAL STATISTICS

Practical 6

Spring 2017

`

- One Sample t test and t Interval
  Two Sample - Unpaired t tests and t Intervals
  Two Sample - Paired t tests and t Intervals

- Handout 2 of 2

In this Practical you will learn to:

1. Construct hypotheses for all of the above tests.
2. Interpret the results of all of the above tests using both CI's and p values.

THERE ARE TWO SECTIONS IN THESE NOTES:

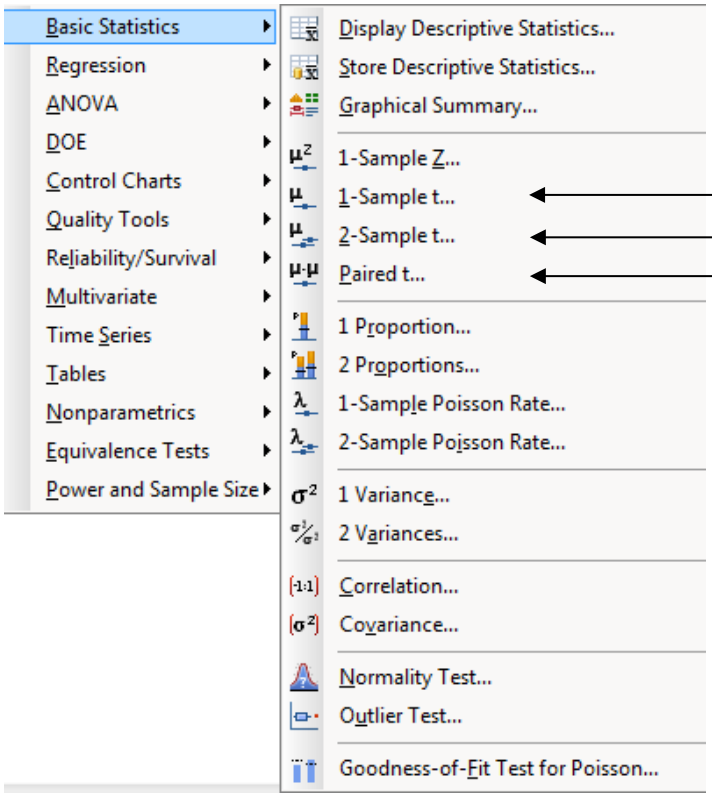
1. INSTRUCTIONS ON HOW TO PERFORM TASKS USING MINITAB.

2. A LIST OF EXERCISES TO DO USING THE ABOVE COMMANDS

C.A. Howie
Computing Science & Maths
Faculty of Natural Sciences
University of Stirling

Introduction

In this Workshop we will be using different options from the following menus:



Sections

1-Sample t... ← 1

2-Sample t... ← 2.1 & 2.2

Paired t... ← 3

Important

If the data can be assumed to be Normally distributed, we will perform parametric tests and construct parametric confidence intervals.

Parametric tests assume that the data is normally distributed and use the mean as the measure of location and standard deviation (variance) as the measure of spread.

All tests and confidence intervals in this Practical are parametric.

1. One Sample t test

Situation : *The data is collected from one group of experimental units*
Question : *About the mean response.*
Condition: **Is the data Normally distributed?** **YES**
Is the variance known for the population? **NO - estimated from sample**

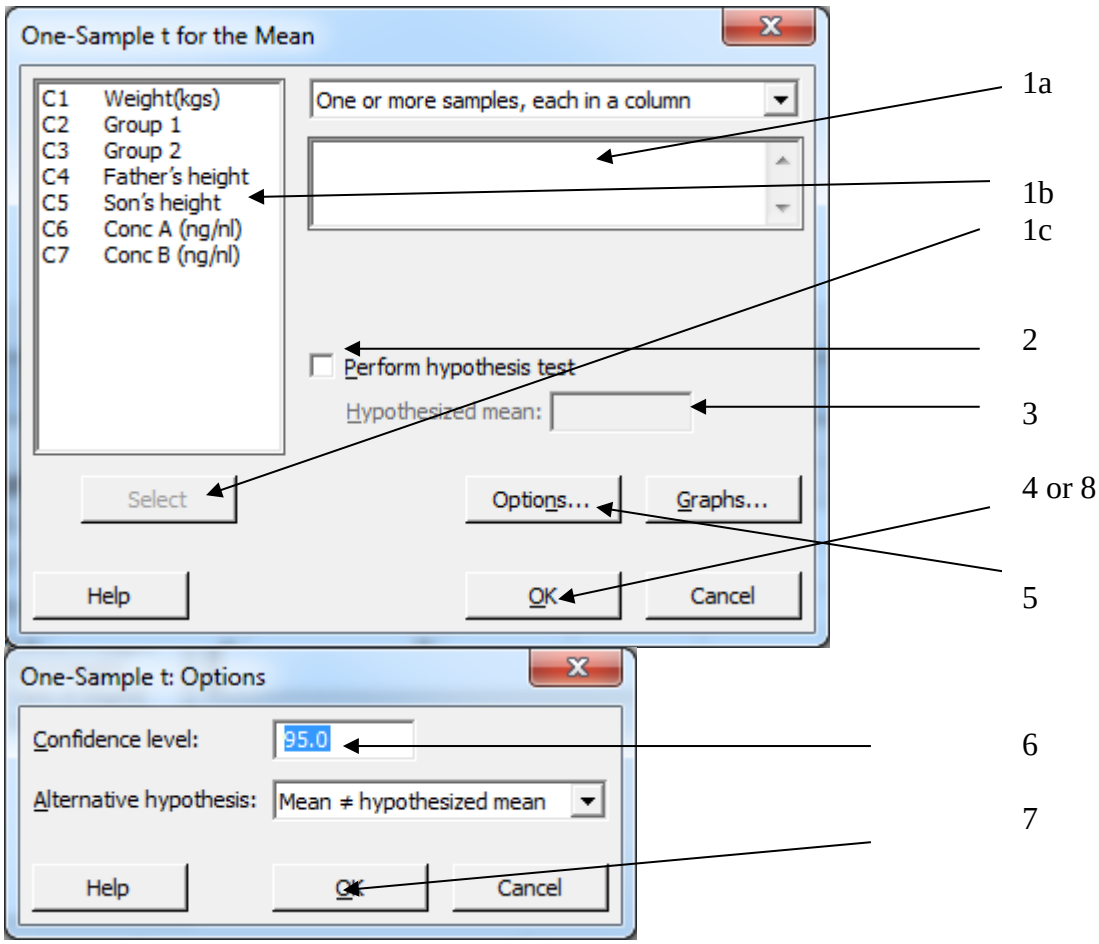
You can perform any of these 3 tests :

Null hypothesis H_0 : $\mu = \mu_o$ against

Alternative H_1 : $\mu \neq \mu_o$ or $\mu > \mu_o$ or $\mu < \mu_o$

Access the **Stat Menu** -> **Basic Statistics** -> **1-Sample t**
In the dialogue box,

1. Select the Columns you want to use as Samples in Columns
   - 1a Click on 1a
   - 1b Choose the column from the list
   - 1c Click Select
2. Choose Perform Hypothesis Test
3. Enter the Hypothesized mean :
4. If you do not want to change the Confidence Level or H_1 – Click OK
5. If you do want to change the Confidence Level or H_1 – Click Options - the second box appears
6. Change Confidence Level or Alternative
7. Click OK
8. Click OK



2. Two Sample Problems I – Independent Groups

Check whether the data for both groups is normally distributed

2.1 Two Sample (Unpaired t test) – Assuming Equal Variances

Situation : *The data is collected from two **independent** groups of experimental units*
Question : *About the mean responses.*
Condition : **Is the data Normally distributed? YES**
Are the variances equal? YES (See Section 5)

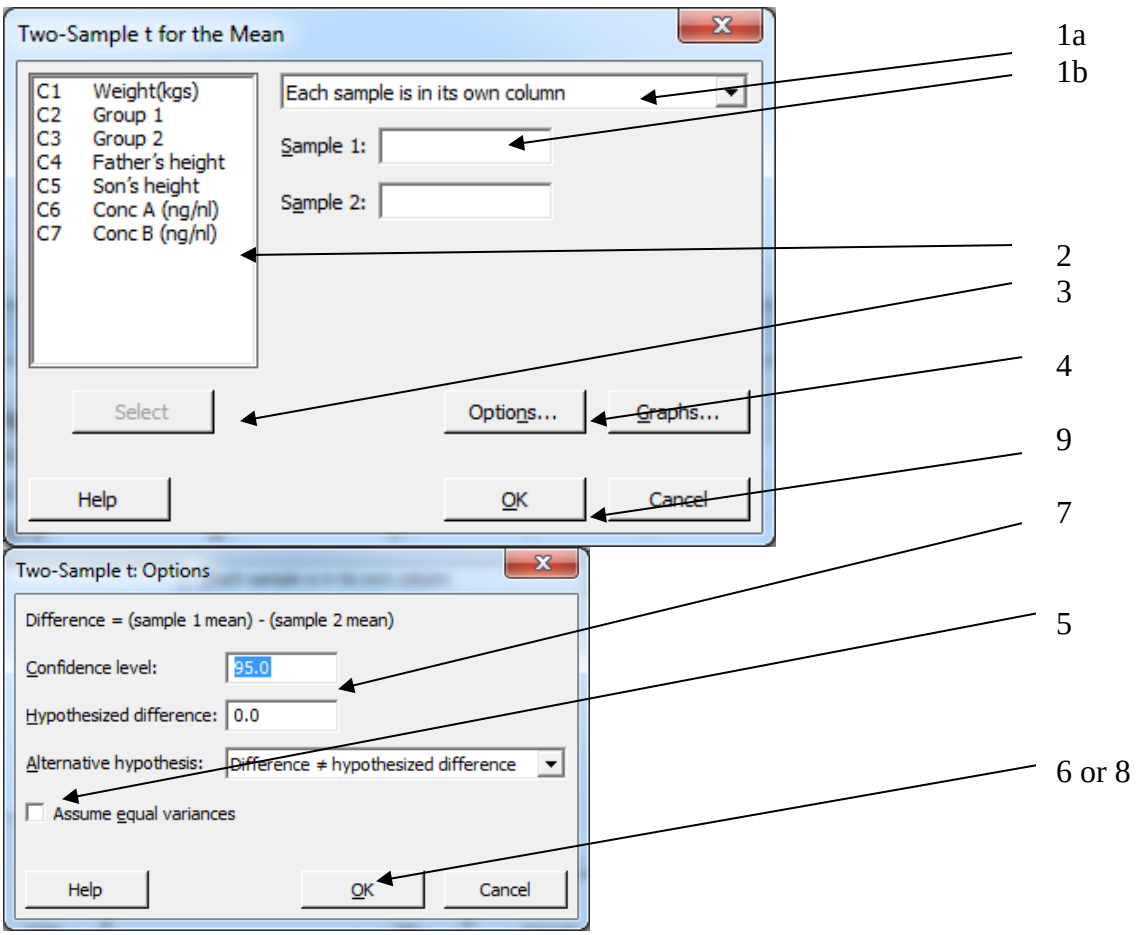
You can perform any of these 3 tests :

Null hypothesis H_0 : $\mu_1 = \mu_2$ against

Alternative H_1 : $\mu_1 \neq \mu_2$ or $\mu_1 > \mu_2$ or $\mu_1 < \mu_2$

Access the **Stat Menu** -> **Basic Statistics** -> **2-Sample t**
In the dialogue box,

1. Choose either Both Samples are in one column or Each sample in its own column depending on your data layout; Click 1b below
2. Choose the columns from the list
3. Click Select
4. To assume equal variance – Choose Options – Second dialog box appears
5. Click Assume equal variance
6. If you do want to change the Confidence Level or H_1– Click OK
7. Change Confidence Level or H_1 or Difference (Mostly leave as 0)
8. Click OK
9. Click OK



2.3 Two Sample (Unpaired t test) – Not assuming Equal Variances

Situation : *The data is collected from two **independent** groups of experimental units*
Question : *About the mean responses.*

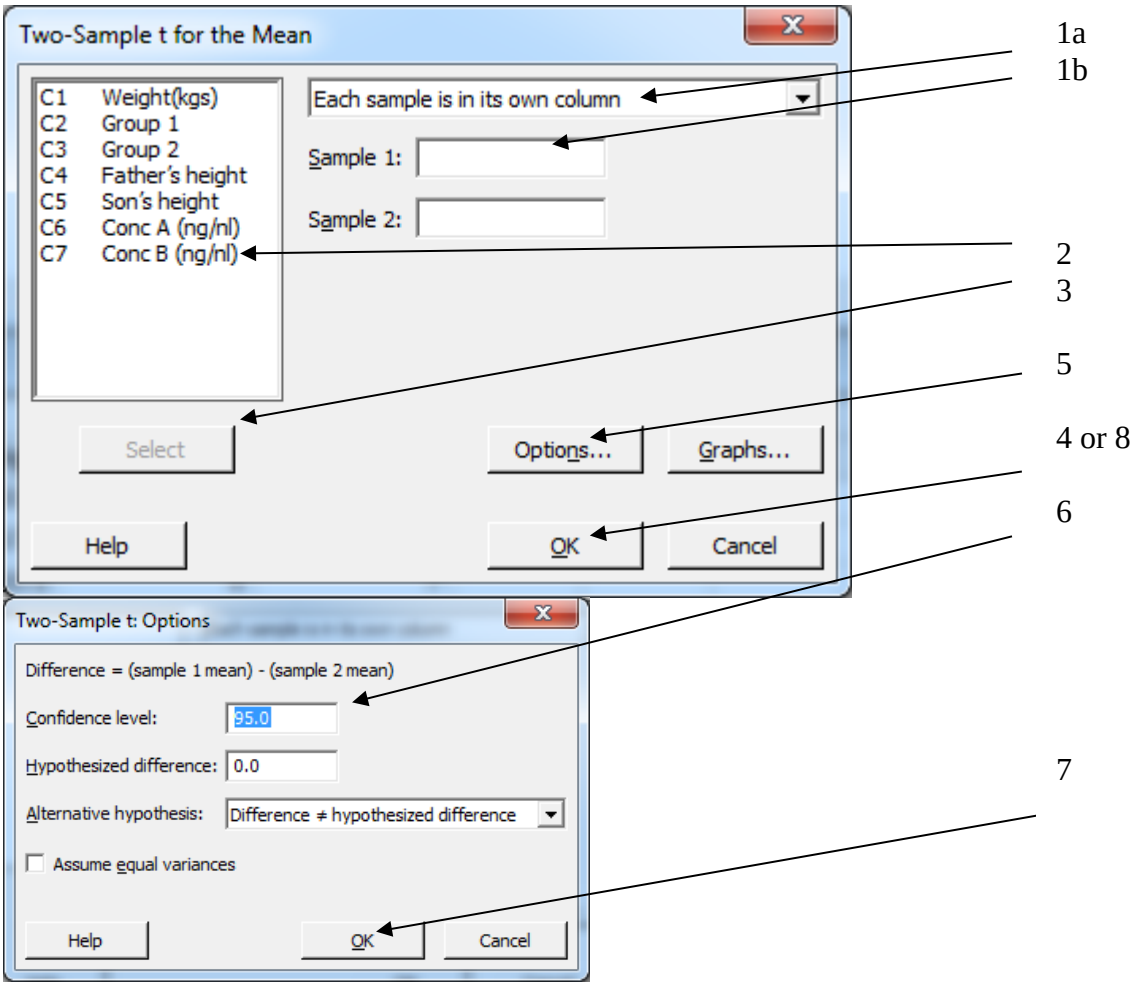
Condition : **Is the data Normally distributed? YES**
Are the variances equal? NO (see next week)

You can perform any of these 3 tests :

Null hypothesis H_0 : $\mu_1 = \mu_2$ against
Alternative H_1 : $\mu_1 \neq \mu_2$ or $\mu_1 > \mu_2$ or $\mu_1 < \mu_2$

Access the **Stat Menu** -> **Basic Statistics** -> **2-Sample t**
In the dialogue box,

1. Choose either Both Samples are in one column or Each sample in its own column depending on your data layout, then Click 1b Below
2. Choose the columns from the list
3. Click Select
4. If you do want to change the Confidence Level or H_1, Click OK
5. If you want to change the Confidence Level or H_1, Choose Options – Second dialog box appears
6. Change Confidence Level or H_1 or Difference (Mostly leave as 0)
7. Click OK
8. Click OK



3. Two Sample Problems II – Paired Data

Situation : *The data is collected from two **related** groups of experimental units.*
e.g. The same group measured twice or is the 'units' are matched.
Question : *About the mean difference in response being zero.*

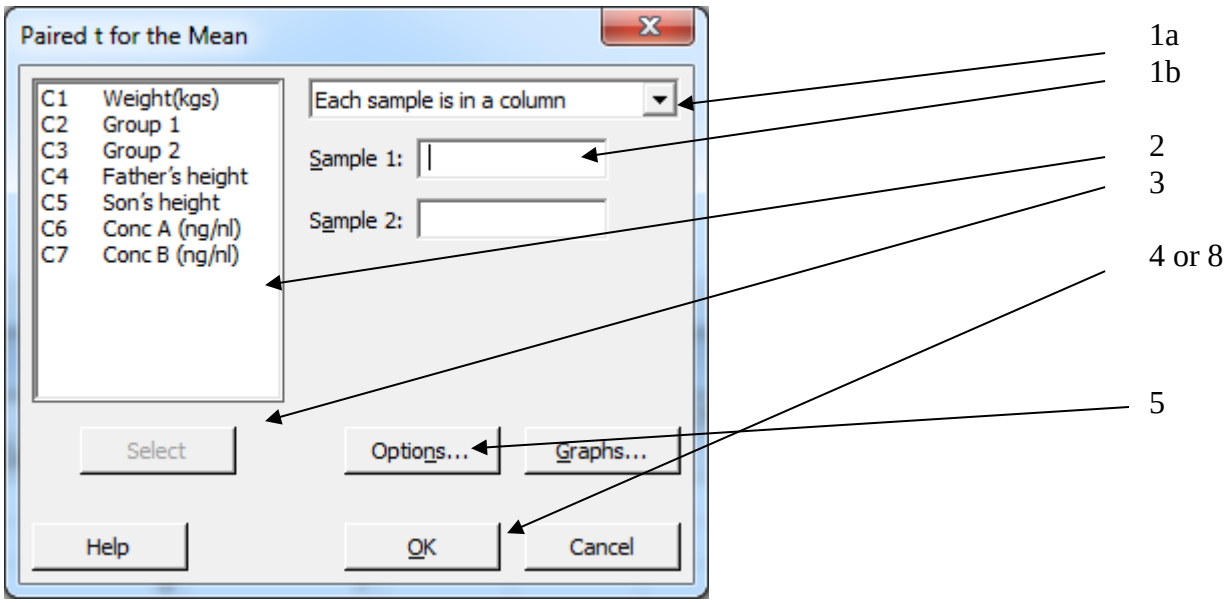
You can perform any of these 3 tests :

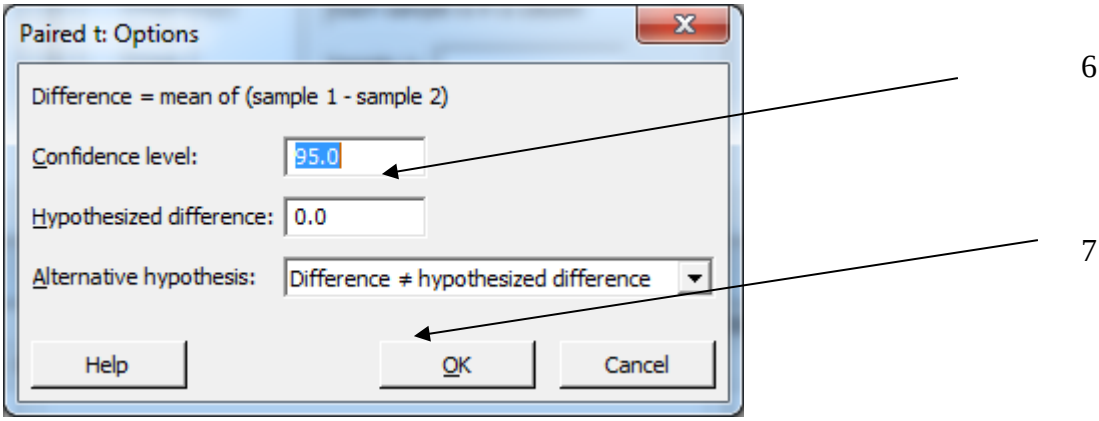
Null hypothesis H_0 : $\mu_d = \mu_0$ (μ_0 is usually 0) against
Alternative H_1 : $\mu \neq \mu_o$ or $\mu > \mu_o$ or $\mu < \mu_o$

Access the Stat Menu -> Basic Statistics -> Paired t

In the dialogue box,

1. Choose Each sample in its own column, then Click 1b Below
2. Choose the columns from the list
3. Click Select
4. If you do not want to change the Confidence Level or H_1 – Click OK
5. If you do want to change the Confidence Level or H_1 – Click Options – second dialog box appears
6. Change Confidence Level or H_1 or Hypothesized Difference (Mostly leave as 0)
7. Click OK
8. Click OK





4. Example Output

All these should two-sided tests – you may have to undertake one-sided ones too!!

4.1 <u>One Sample t test</u>

One-Sample T: Weight(kgs)

```
Test of μ = 3.6 vs ≠ 3.6

Variable      N    Mean   StDev  SE Mean        95% CI          T      P
Weight(kgs)  36  3.2861  0.5514   0.0919  (3.0995, 3.4727)  -3.42  0.002
```

Hypotheses: equality vs inequality

CI : states level

Observed Test Stat

p value < 0.05 so can reject H_o for H_1 at 5% level in this case

4.2 <u>Unpaired t test (assuming equal variances)</u>

Two-Sample T-Test and CI: Group 1, Group 2

```
Two-sample T for Group 1 vs Group 2

          N   Mean  StDev  SE Mean
Group 1  10  3.620  0.294    0.093
Group 2  10  3.250  0.259    0.082

Difference = μ (Group 1) - μ (Group 2)
Estimate for difference:  0.370
95% CI for difference:  (0.110, 0.630)
T-Test of difference = 0 (vs ≠): T-Value = 2.99  P-Value = 0.008  DF = 18
Both use Pooled StDev = 0.2770
```

CI : states level

Observed Test Stat

Pooled sd : so this test assumes equal variance

Hypotheses: equality vs inequality

p value if < 0.05 sig at 5% level i.e. reject H_o for H_1 in this case

4.3 <u>Unpaired t test (not assuming equal variances)</u>

Two-Sample T-Test and CI: Group 1, Group 2

```
Two-sample T for Group 1 vs Group 2

   N   Mean  StDev  SE Mean
Group 1  10  3.620  0.294    0.093
Group 2  10  3.250  0.259    0.082

Difference = μ (Group 1) - μ (Group 2)
Estimate for difference:  0.370
95% CI for difference:  (0.109, 0.631)
T-Test of difference = 0 (vs ≠): T-Value = 2.99  P-Value = 0.008  DF = 17
```

CI : states level

Observed Test Stat

Hypotheses: in this case H_o : means equal, H_1 means different

p value if < 0.05 sig at 5% level i.e. reject H_o for H_1 in this case

4.4 Paired t test

Paired T-Test and CI: Father's height, Son's height

```
Paired T for Father's height - Son's height

                N   Mean  StDev  SE Mean
Father's height 8  68.59   3.65     1.29
Son's height    8  67.69   2.95     1.04
Difference      8   0.90   3.11     1.10

95% CI for mean difference: (-1.70, 3.50)
T-Test of mean difference = 0 (vs ≠ 0): T-Value = 0.82  P-Value = 0.440
```

CI : states level

Observed Test Stat

Hypotheses : equality vs inequality

p value if < 0.05 sig at 5% level i.e. cannot reject H_0 for H_1 in this case

EXERCISES

1. Perform the appropriate statistical test to formally answer the questions in this Practical.

   In each case,

   (i) you should verify that the data is Normally distributed.

   (ii) whether the test is one sample or two sample?

   (iii) which test is appropriate?

   (iv) whether the test is :
   (a) one tailed - if the question specifies a direction
   (b) two tailed - if the question asks if there is a difference

   (v) interpret the output

   For Example:

. (a) Sufficient evidence of a difference at the given Confidence Level (usually 95%)
   or equivalently Significance Level (5%)
   i.e. **Statistically significant difference**
   : if the Confidence Interval does not include zero
   : the p value will be less than the significance level

   (b) Insufficient evidence at the given Confidence Level (usually 95%)
   or equivalently Significance Level (5%)
   i.e. **Non-significant difference**
   : if the Confidence Interval does includes zero
   : the p value will be greater than the significance level

3. Do your answers tie in with your subjective impressions?

Questions

By the end of this practical : Make sure that you are confident with :

(i) Getting appropriate graphs and numerical summaries
(ii) Drawing subjective impressions
(iii) Checking whether data is normally distributed
(iv) All the new test and confidence intervals from this practical

1. The birth weights (kgs) of 36 babies born after normal pregnancies of 40 weeks were :

3.5	4.1	2.8	3.2	2.8	3.1	3.4	3.0	2.3
3.8	2.7	3.7	3.9	2.6	2.7	3.1	2.2	2.9
3.2	3.7	3.3	4.3	3.4	3.5	4.6	3.1	3.4
3.5	3.5	3.8	2.4	3.0	3.6	4.0	2.9	3.3

Is the mean birth weight different from 3.6kgs?

2. Drug levels (in ng/ml) in blood samples from two groups of subjects gave values of :

Group 1	:	3.3	3.7	3.5	4.1	3.4	3.5	4.0
		3.8	3.2	3.7				
Group 2	:	3.2	3.6	3.1	3.4	3.0	3.4	2.8
		3.1	3.3	3.6				

Is there a difference in the means?

3. In an experiment to detect any relationship between the heights of Aberdonian fathers and their eldest sons, eight pairs of fathers and sons were selected at random from the city population and their exact heights recorded, in inches, as follows :

Father's height	:	63.1	74.5	70.4	72.1	65.9	67.1	68.8	66.8
Son's height	:	65.6	67.9	69.6	70.8	64.5	68.2	71.5	63.4

Is there a difference in height between fathers and sons?

4. Concentrations of two drugs A and B were measured in the same subjects following equal oral doses of a drug on different days. The pairs (A,B) concentrations are

(9,9) (8,8) (4,7) (4,8) (3,5) (3,6) (3,7) (1,6) (2,2) ng/ml after 1 hour

Does B result in higher concentrations than A?

5. Use Minitab to answer today's By Hand questions too!!