17.1 Variation

YOUR NOTES ⬇

CONTENTS

17.1.1 VARIATION: PHENOTYPE

Variation: Phenotype

- The observable characteristics of an organism are its phenotype
- Phenotypic variation is the **difference in phenotypes** between organisms of the same **species**
- In some cases, **phenotypic variation** is explained by **genetic** factors
    - For example, the four different blood groups observed in human populations are due to different individuals within the population having two of **three** possible **alleles** for the single ABO gene
- In other cases, **phenotypic variation** is explained by **environmental** factors
    - For example, clones of plants with exactly the same genetic information (DNA) will grow to different heights when grown in different environmental conditions
- Phenotypic variation can also be explained by a **combination of genetic and environmental factors**
    - For example, the recessive allele that causes sickle cell anaemia has a high frequency in populations where malaria is prevalent due to heterozygous individuals being resistant to malaria
- The complete phenotype of an organism is determined by the expression of its genotype and the interaction of the environment on this:

Phenotype = Genotype + Environment

17.1 Variation

YOUR NOTES ⬇

Genetic variation

- Organisms of the same species will have very similar genotypes, but two individuals (even twins) will have differences between their DNA base sequences
- Considering the size of genomes, these differences are small between individuals of the same species
- The small differences in DNA base sequences between individual organisms within a species population is called **genetic variation**
- Genetic variation is **transferred** from one generation to the next and it **generates phenotypic variation** within a species population
- Genetic variation is caused by the following processes as they result in a **new combination of alleles** in a gamete or individual:
  - **Independent assortment** of homologous chromosomes during metaphase I
  - **Crossing over** of non-sister chromatids during prophase I
  - **Random fusion** of gametes during fertilization
- Mutation results in the **generation of new alleles**
  - The new allele may be advantageous, disadvantageous or have no apparent effect on phenotype (due to the fact that the genetic code is **degenerate**
  - New alleles are not always seen in the individual that they first occur in
  - They can remain hidden (not expressed) within a population for several generations before they contribute to phenotypic variation
- Genes can have **varying effects** on an organism's phenotype
  - The phenotype may be affected by a single gene or by several
  - The effect that the gene has on the phenotype may be large, small and/or additive

17.1 Variation

YOUR NOTES ⬇

Sources of genetic variation table

Process	Mechanism	Consequences
Independent assortment of homologous chromosomes during metaphase I	Random alignment of chromosomes results in different combinations of chromosomes and different allele combinations in each gamete	Genetic variation between gametes produced by an individual
Crossing over of non-sister chromatids during prophase I	Exchange of genetic material between non-sister chromatids leads to new combinations of alleles on chromosomes. It can also break linkage between genes	Genetic variation between gametes produced by an individual
Random fusion of gametes during fertilization	Any male gamete is able to fuse with any female gamete (Random mating in a species population)	Genetic variation between zygotes and resulting individuals
Mutation	Random change in the DNA base sequence results in the generation of a new allele. Mutation must exist within gametes for it to passed onto future generations	Genetic variation between individuals within a species population

Environmental factors

- The **environment** that an organism lives in can also have an impact on its phenotype
- Different environments around the globe experience very different conditions in terms of the:
  - Length of sunlight hours (which may be seasonal)
  - Supply of nutrients (food)
  - Availability of water
  - Temperature range
  - Oxygen levels
- Changes in the factors above can affect how organisms **grow and develop**
  - For example, plants with a tall genotype growing in an environment that is depleted in minerals, sunlight and water will not be able to grow to their full potential size determined by genetics
- Variation in phenotype caused solely by environmental pressures or factors cannot be **inherited** by an organism's offspring
  - Only alterations to the genetic component of gametes will ever be inherited

17.1 Variation

YOUR NOTES ⬇

Exam Tip

Some questions in the exam may ask you to explain why the variation in phenotype due to genetics is inherited but the variation in phenotype due to environmental factors is not. This is because genetic variation directly affects the DNA of the gametes but variation in phenotype caused by the environment does not.

YOUR NOTES ⬇

17.1 Variation

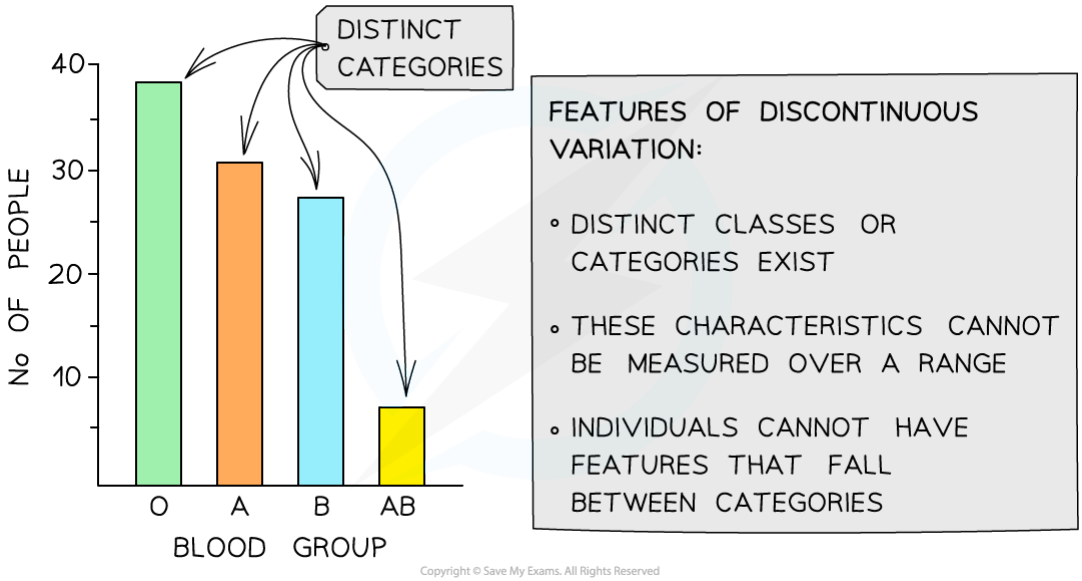
17.1.2 VARIATION: DISCONTINUOUS & CONTINUOUS

Variation: Discontinuous & Continuous

- The term **variation** refers to the differences that exist between at least two things (be it a level, amount, quantity or feature of something)
- In relation to natural selection, variation refers to the **differences that exist between individuals of a species**
  - This may also be referred to as **intraspecific** variation
- Variation observed in the **phenotypes** of organisms can be due to qualitative or quantitative differences

Discontinuous variation

- **Qualitative differences** in the phenotypes of individuals within a population give rise to **discontinuous variation**
- Qualitative differences fall into discrete and distinguishable **categories,** usually with no intermediates (a feature can't fall in between categories)
  - For example, there are four possible ABO blood groups in humans; a person can only have one of them
- It is easy to identify discontinuous variation when it is present in a table or graph due to the distinct categories that exist when data is plotted for particular characteristics



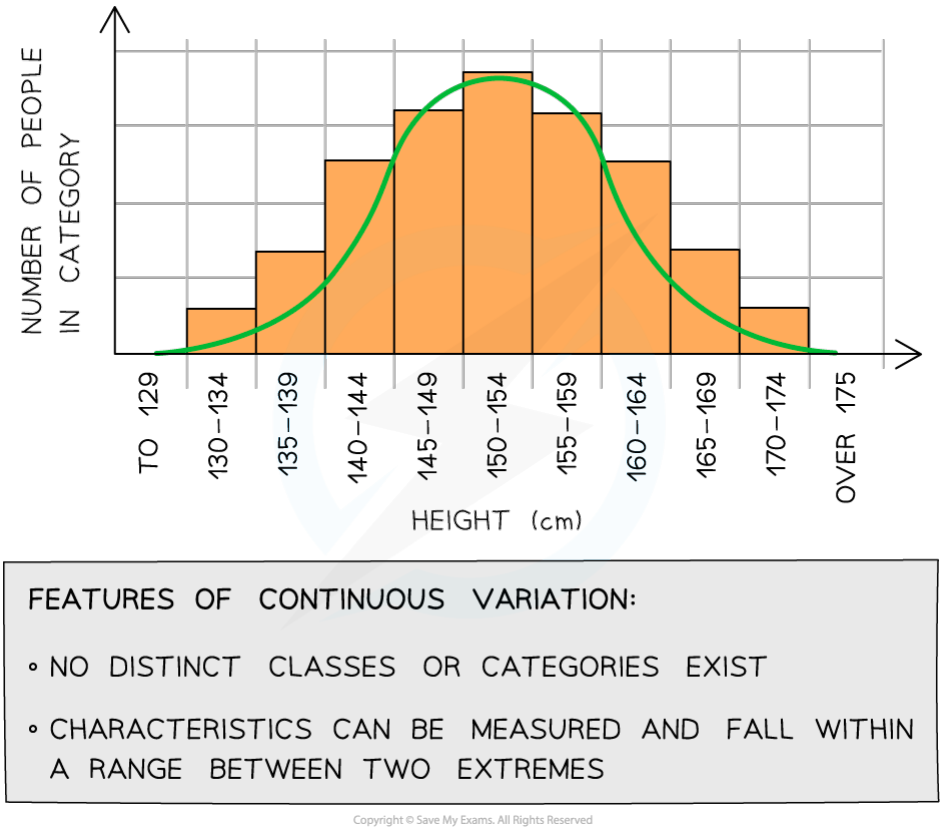
Graph showing population variation in blood types: an example of discontinuous variation with qualitative differences

17.1 Variation

YOUR NOTES ⬇

Continuous variation

- Continuous variation occurs when there are **quantitative differences** in the phenotypes of individuals within a population for particular characteristics
- Quantitative differences do not fall into discrete categories like in discontinuous variation
- Instead for these features, a **range of values** exist between two extremes within which the phenotype will fall
    - For example, the mass or height of a human is an example of continuous variation
- The lack of categories and the presence of a range of values can be used to identify continuous variation when it is presented in a table or graph



Graph showing population variation in height: an example of continuous variation with quantitative differences

17.1 Variation

YOUR NOTES ⬇

The Genetic Basis of Variation

- **Discontinuous variation** refers to the differences between individuals of a species where the differences are **qualitative** (categoric)
- **Continuous variation** is the differences between individuals of a species where the differences are **quantitative** (measurable)
- Each type of variation can be explained by **genetic** and / or **environmental factors**

Genetic basis of discontinuous variation

- This type of variation occurs solely due to **genetic factors**
- The environment has no direct effect
  - Phenotype = **genotype**
- At the genetic level:
  - Different **genes** have **different effects** on the phenotype
  - Different **alleles** at a single gene locus have a **large effect** on the phenotype
  - Remember diploid organisms will inherit two alleles of each gene, these alleles can be the same or different
- A good example of this is the *F8* gene that codes for the blood-clotting protein Factor VIII
  - The different alleles at the *F8* gene locus dictate whether or not normal Factor VIII is produced and whether the individual has the condition haemophilia

Genetic basis of continuous variation

- This type of variation is caused by an **interaction between genetics and the environment**
- Phenotype = **genotype + environment**
- At the genetic level:
  - Different **alleles** at a single locus have a **small effect** on the phenotype
  - Different **genes** can have the **same effect** on the phenotype and these add together to have an **additive effect**
  - If a large number of genes have a combined effect on the phenotype they are known as **polygenes**

17.1 Variation

YOUR NOTES ⬇

The additive effect of genes

- The height of a plant is controlled by two unlinked genes **H** / **h** and **T** / **t**
- The two genes have an **additive** effect
- The recessive alleles **h** and **t** contribute ***x*** cm to the plants height
- The dominant alleles **H** and **T** contribute 2***x*** cm to the plants height
- The following genotypes will have the following phenotypes:
  - **h h t t :** ***x*** + ***x*** + ***x*** + ***x*** = 4***x*** cm
  - **H H T T :** 2***x*** + 2***x*** + 2***x*** + 2***x*** = 8***x*** cm
  - **H h T t :** 2***x*** + ***x*** + 2***x*** + ***x*** = 6***x*** cm
  - **H H T t :** 2***x*** + 2***x*** + 2***x*** + ***x*** = 7***x*** cm
  - **H h T T :** 2***x*** + ***x*** + 2***x*** + 2***x*** = 7***x*** cm
  - **h h T t :** ***x*** + ***x*** + 2***x*** + ***x*** = 5***x*** cm
  - **H h t t :** 2***x*** + ***x*** + ***x*** + ***x*** = 5***x*** cm

Exam Tip

Be careful when answering questions that involve polygenes or genes with an additive effect. It is not a given that each gene will have the same effect on the phenotype as in the example above so make sure to double check the information you have been given in the question.

17.1 Variation

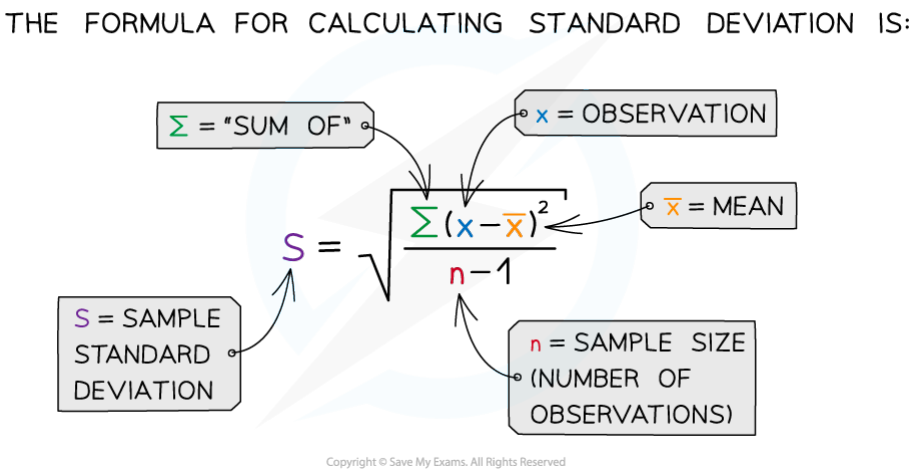
YOUR NOTES ⬇

17.1.3 VARIATION: T-TEST METHOD

Variation: t-test Method

- A statistical test called the **t-test** can be used to **compare the means of two sets of data** and determine whether they are **significantly** different or not
  - The formula for the t-test will be provided in the exam, but formulae for how to calculate the number of **degrees of freedom** is not provided in the exam and must be learnt
- The sets of data must follow a rough **normal distribution**, be **continuous** and the **standard deviations** should be approximately equal
- The standard deviation (*s*) must be calculated for each data set before the t-test can be carried out
- A **null hypothesis** should also be given
  - This is a statement of what we would expect if there is **no significant difference** between two means, and that any differences seen are due to **chance**
- If there is a statistically significant difference between the means of two sets of data, then the observation is not down to chance and the **null hypothesis** can be **rejected**

Calculating the standard deviation



Using the t-test to compare two means

- The steps below outline the general steps in a *t* test; for a worked example see the next page
- Null hypothesis: there is no statistically significant difference between the means of sample 1 and sample 2
- **Step 1**: Calculate the mean for each data set:

$\bar{x}_1$ = the mean for sample 1, and $\bar{x}_2$ = the mean for sample 2

17.1 Variation

YOUR NOTES ⬇

- **Step 2**: Calculate the **standard deviation** for each set of data, s_1 = standard deviation of sample 1 and s_2 = standard deviation of sample 2

$$s_n = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n - 1}}$$

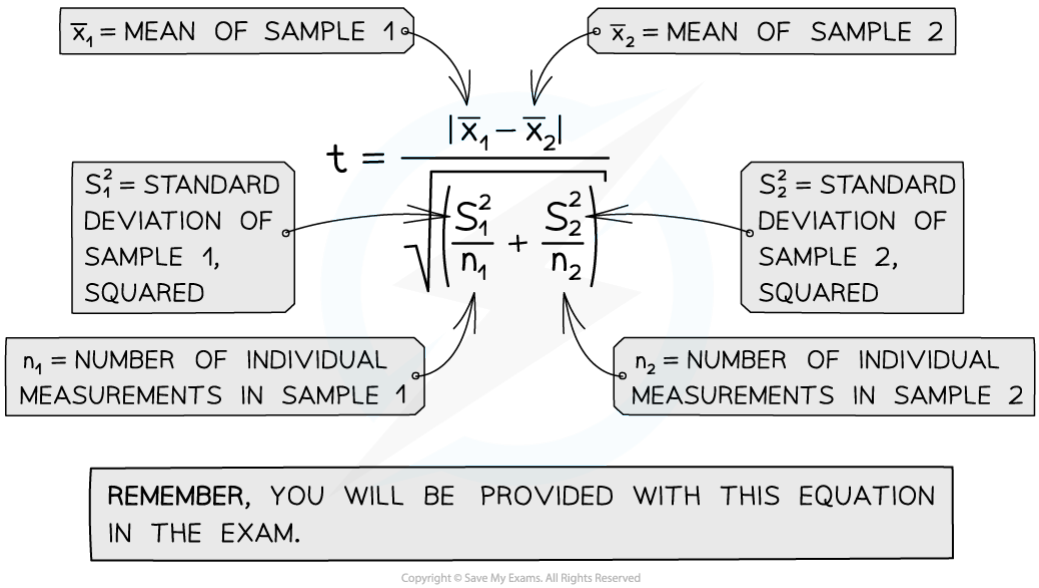
- **Step 3**: Square the standard deviation and divide by *n* (the number of observations) in each sample, for both samples:

$$\frac{s_1^2}{n_1} \text{ and } \frac{s_2^2}{n_2}$$

- **Step 4:** Add the values from step 3 together and take the square root:

$$\sqrt{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})}$$

- **Step 5:** Divide the difference between the two means (see step 1) with the value calculated in step 4 to get the *t* value:



17.1 Variation

YOUR NOTES ⬇

- **Step 6:** Calculate the **degrees of freedom** (v) for the whole data set (remember the formulae for this will **not** be given in the exam):

$$v = (n_1 - 1) + (n_2 - 1)$$

- **Step 7:** Look at a table that relates t values to the probability that the differences between data sets is due to chance to find where the t value for the degrees of freedom (v) calculated lies

17.1 Variation

YOUR NOTES ⬇

T values table

Degrees of freedom	Value of t			
1	6.31	12.7	63.7	63.6
2	2.92	4.30	9.93	31.6
3	2.35	3.18	5.84	12.9
4	2.13	2.78	4.60	8.61
5	2.02	2.57	4.03	6.87
6	1.94	2.45	3.71	5.96
7	1.90	2.37	3.50	5.41
8	1.86	2.31	3.36	5.04
9	1.83	2.26	3.25	4.78
10	1.81	2.23	3.17	4.59
Probability that chance could have produced this value of t	0.10	0.05	0.01	0.001
Confidence level	10%	5%	1%	0.1%

17.1 Variation

YOUR NOTES ⬇

- **Step 8**: The greater the *t* value calculated (for any degree of freedom), the lower the probability of chance causing any significant difference between the two sample means
    - Identify where the *t* value calculated lies with respect to the confidence levels provided
    - If the *t* value is **greater than the critical value** (obtained from the table at the critical probability of 0.05) then any difference between the two data sets is less likely to be **due to chance,** so the **null hypothesis can be rejected**
    - If the *t* value is less than the critical value given at a confidence of 5% / the probability that any difference is down to chance is above 0.05; then an assumption can be made that the differences between the means of the two sets of data are not significant and the **null hypothesis is accepted**
- Using the table above, if a value of *t* was calculated to be 2.38 at 5 degrees of freedom, then it lies between 2.02 and 2.57, so the probability that chance produced any differences between the two means is between 10% and 5%; the null hypothesis would be accepted in this situation

Exam Tip

If you need to calculate the t value you will be given the formula in the exam. Generally questions on the t-test require you to:

• Know why a *t*-test is being used to analyse the data

• State the null hypothesis

• Know how the degrees of freedom was calculated

• State the conclusion (are the differences between the two means significant or not)

17.1 Variation

YOUR NOTES ⬇

17.1.4 VARIATION: T-TEST WORKED EXAMPLE

Variation: t-test Worked Example

Worked example: T-test

The ear length of two populations of rabbits was measured.

Ear lengths of population A (mm):
62, 60, 59, 61, 60, 58, 59, 60, 57, 56, 59, 58, 60, 59, 57

Ear lengths of population B (mm):
58, 59, 57, 59, 59, 57, 55, 60, 57, 58, 59, 58, 57, 58, 59

Use the t-test to determine whether there is a significant difference in ear length between the two populations.

Solution

- Null hypothesis: There is **no *significant* difference** between the ear lengths of the rabbits in populations *A* and *B*
- Sample sizes:
  - Population *A*: $n_1 = 15$
  - Population *B*: $n_2 = 15$
- **Step 1:** Calculate the **mean** for each data set:
  - Mean for population A $\bar{x}_1 = 885 / 15 = 59$ mm
  - Mean for population B $\bar{x}_2 = 870 / 15 = 58$ mm
- **Step 2:** Calculate the **standard deviation** (*s*) for each set of data:

Worked example t-test table 1

17.1 Variation

YOUR NOTES ⬇

Population A		Population B	
Difference between value and mean $(x - \bar{x})$	Difference between value and mean squared $(x - \bar{x})^2$	Difference between value and mean $(x - \bar{x})$	Difference between value and mean squared $(x - \bar{x})^2$
62 – 59 = 3	9	58 – 58 = 0	0
60 – 59 = 1	1	59 – 58 = 1	1
59 – 59 = 0	0	57 – 58 = –1	1
61 – 59 = 2	4	59 – 58 = 1	1
60 – 59 = 1	1	59 – 58 = 1	1
58 – 59 = –1	1	57 – 58 = –1	1
59 – 59 = 0	0	55 – 58 = –3	9
60 – 59 = 1	1	60 – 58 = 2	4
57 – 59 = –2	4	57 – 58 = –1	1
56 – 59 = –3	9	58 – 58 = 0	0
59 – 59 = 0	0	59 – 58 = 1	1
58 – 59 = –1	1	58 – 58 = 0	0
60 – 59 = 1	1	57 – 58 = –1	1
59 – 59 = 0	0	58 – 58 = 0	0
57 – 59 = –2	4	59 – 58 = 1	1
Total $\Sigma(x - \bar{x})^2$	36	Total $\Sigma(x - \bar{x})^2$	22

17.1 Variation

YOUR NOTES ⬇

- Divide the sum of each square by *n* – 1 for each data set, and take the square root of each value:

Worked example t-test table 2

Population A ($n_1 = 15$)	Population B ($n_2 = 15$)
$n_1 - 1 = 14$	$n_2 - 1 = 14$
$\Sigma(x-\bar{x})^2 = 36$ so $36/14 = 2.57$	$\Sigma(x-\bar{x})^2 = 22$ so $22/14 = 1.57$
$\sqrt{2.57} = 1.60$	$\sqrt{1.57} = 1.25$
$s_1 = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}} = s_1 = 1.60$	$s_2 = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}} = s_2 = 1.25$

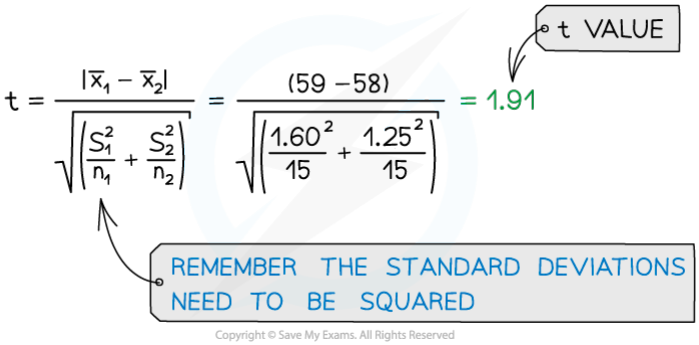
- **Step 3 to 5**: Sub all known values into the *t*-test equation by:
  - **Step 3**: Square the standard deviation and divide by *n* (the number of observations) in each sample, for both samples:
  - **Step 4**: Add the values from step 3 together and find the square root
  - **Step 5**: Divide the difference between the two means by the value from step 4

17.1 Variation

YOUR NOTES ⬇

Worked example t-test table 3

Population A	Population B
$\bar{x}_1 = 59$	$\bar{x}_2 = 58$
$s_1 = 1.60$	$s_2 = 1.25$
$n_1 = 15$	$n_2 = 15$



$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)}} = \frac{(59 - 58)}{\sqrt{\left(\frac{1.60^2}{15} + \frac{1.25^2}{15}\right)}} = 1.91$$

- **Step 6:** Calculate the **degrees of freedom** *(v)* for all the data:
  - $v = (n_1 - 1) + (n_2 - 1) = 14 + 14 = 28$
- **Step 7:** Look at a table that relates *t* values to the probability that the differences between data sets is due to chance to find where the *t* value of 1.91 for 28 degrees of freedom (*v*) calculated lies

17.1 Variation

YOUR NOTES ⬇

T value worked example table

Degrees of freedom	Value of t			
28	1.70	2.05	2.76	3.67
Probability that chance could have produced this value of t	0.1	0.05	0.01	0.001

- **Step 8:** Draw a conclusion about the statistical relevance of the data:
  - A *t* value of 1.91 represents a probability between 0.05 and 0.1 which is greater than the critical value of 0.05.
  - This means the null hypothesis should be **accepted**, as there are **no significant differences** between the two sets of results (any differences between the means of the ear length of rabbits in the two populations are due to chance)