

Quantitative Genetics

- Mendelian traits were qualitative in nature and were classified into distinct phenotype and are under the control of only one or a very few genes.
- In contrast, many traits do not fit into separate phenotypic classes (**discontinuous variability**) but form a spectrum of phenotypes that blend from one type to another (**continuous variability**) and is called **quantitative or complex traits or metric traits**.
- E.g. body weight gains, mature plant height, egg or milk production records, yield of grains per acre.
- It is governed by **number of genes** (may be 10-100) and they affect the **degree** to which the phenotypes can be modified by environmental factors.
- All genes act in concert, and also referred as **Polygenic traits**. The genes responsible for phenotypes associated with quantitative traits are called **Quantitative Trait Loci (QTLs)**.

Quantitative Genetics

- It arise from two phenomenon:
- **First**, many are polygenic, they are influenced by genes at many loci. If many loci take part, many genotypes are possible, each producing a slightly different phenotypes.
- **Second**, they arise, when environmental factors affect the phenotype, because environmental differences result in a single genotype producing a range of phenotypes.

Qualitative vs Quantitative Genetics

.Relation between genotype and phenotype:

- **For many discontinuous (qualitative) traits**, the relation between genotype and phenotype is straightforward. Each genotype produces a single phenotype. Dominance and epistasis may allow two or three genotypes to produce the same phenotype but the relation remain simple
 - **For Quantitative traits**, the relation between genotype and phenotype is more complex. If the trait is polygenic, many different genotypes are possible, several of which may produce the same phenotype.
- For e.g., Plant height is determined by three genes at three Loci- A, B, C, each of which has two alleles. **One allele** at each locus (A^+ , B^+ , C^+) encodes a plant hormone that causes the plant to grow 1 cm in height above its baseline height of 10 cm. **The other Allele** at each locus (A^- , B^- , C^-) does not encode plant hormone and hence does not contribute to additional height.
- If all three loci are taken into account, there are a total of $3^3 = 27$ possible multilocus **genotypes** but they all will produce **only 7 phenotypes** (10cm, 11cm, 12cm, 13cm, 14cm, 15cm, 16cm). Some of the genotypes produce the same phenotypes

Hypothetical example of plant height determined by pairs of alleles at each of three loci

Plant Genotype	Doses of Hormone	Height (cm)
$A^-A^- B^-B^- C^-C^-$	0	10
$A^+A^- B^-B^- C^-C^-$	1	11
$A^-A^- B^+B^- C^-C^-$		
$A^-A^- B^-B^- C^-C^+$		
$A^+A^+ B^-B^- C^-C^-$	2	12
$A^-A^- B^+B^+ C^-C^-$		
$A^-A^- B^-B^- C^+C^+$		
$A^+A^- B^+B^- C^-C^-$		
$A^+A^- B^-B^- C^+C^-$		
$A^-A^- B^+B^- C^+C^-$		
$A^+A^+ B^+B^- C^-C^-$	3	13
$A^+A^+ B^-B^- C^+C^-$		
$A^+A^- B^+B^+ C^-C^-$		
$A^-A^- B^+B^+ C^+C^-$		
$A^+A^- B^-B^- C^+C^+$		
$A^-A^- B^+B^- C^+C^+$		
$A^+A^- B^+B^- C^+C^-$		
$A^+A^+ B^+B^+ C^-C^-$	4	14
$A^+A^+ B^+B^- C^+C^-$		
$A^+A^- B^+B^+ C^-C^-$		
$A^-A^- B^+B^+ C^+C^+$		
$A^+A^+ B^-B^- C^+C^+$		
$A^+A^- B^+B^- C^+C^+$		
$A^+A^+ B^+B^+ C^+C^-$	5	15
$A^+A^- B^+B^+ C^+C^+$		
$A^+A^+ B^+B^- C^+C^+$		
$A^+A^+ B^+B^+ C^+C^+$	6	16

Note: Each + allele contributes 1 cm in height above a baseline of 10 cm.

Quantitative Genetics

Relation between genotype and phenotype:

Types of Quantitative Genetics

1. Meristic Characteristics: are measured in whole numbers. It has limited number of distinct phenotypes but the determination of traits is still be continuous.

E.g. a female mouse may have 4, 5 or 6 pups but never 4.13.

2. Threshold Characteristics: which .is simply present or absent. Although it exhibit only two phenotypes, they are considered quantitative because they too are determined by multiple genes and environmental factors.

e.g. presence of some disease, like COVID-19

- ✓ The expression of characteristics depends on an underlying susceptibility (usually referred as liability or risk) that varies continuously.
- ✓ When susceptibility is larger than threshold value, specific trait is expressed.
- ✓ If enough of the susceptibility factors are present, the disease develops, otherwise, it is absent.

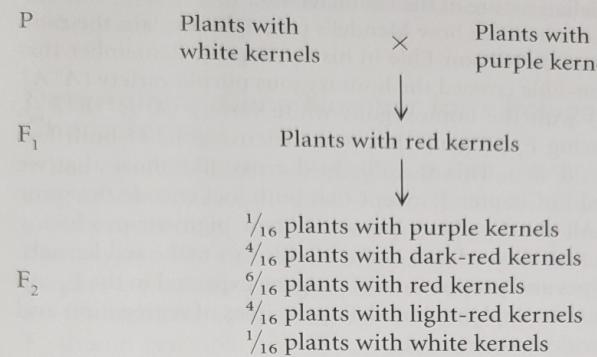
POLYGENIC INHERITANCE

Kernel Color in Wheat

- Studied by Nelsson-Ehle who found that in wheat kernel, the intensity of Red pigmentation was determined by two unlinked loci, each of which had two alleles.
- He crossed a variety of wheat that possessed white kernels with a variety that possessed purple (very dark red).
- Purple and white are encoded by single genotypes but other phenotypes may result from different genotypes.

POLYGENIC INHERITANCE

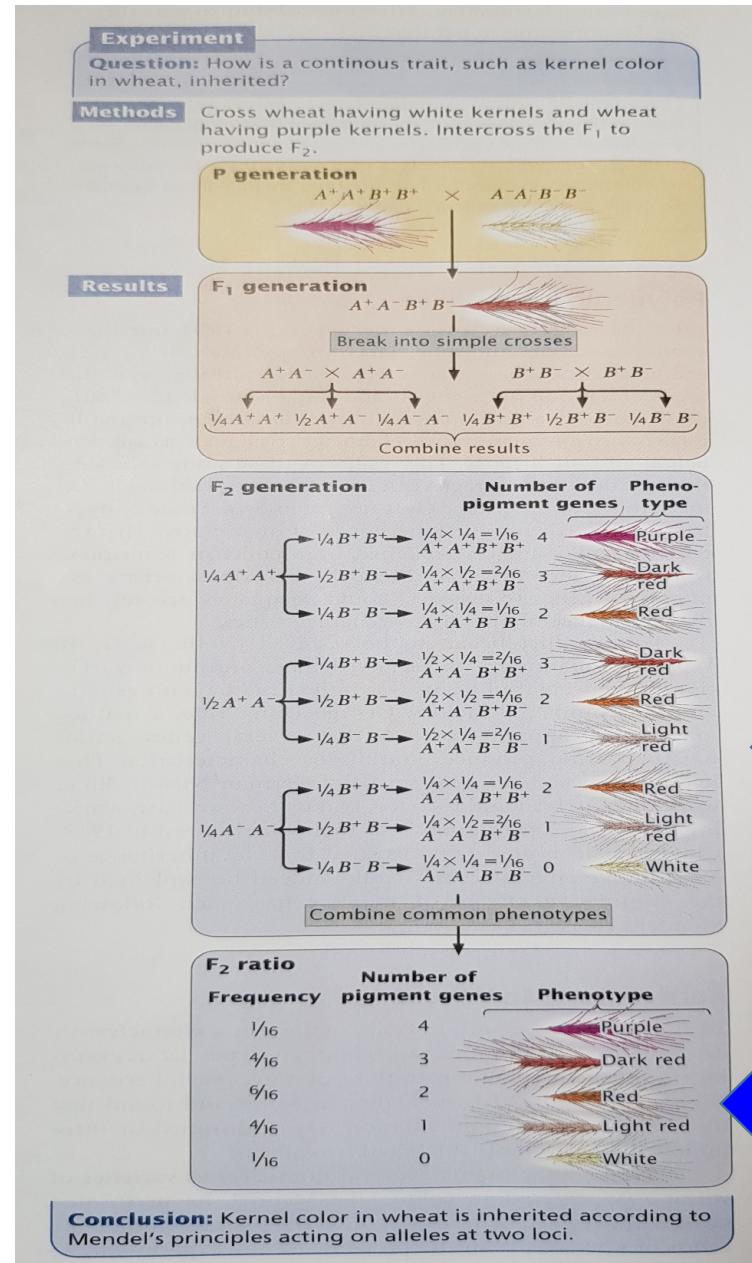
Kernel Color in Wheat



Nilsson-Ehle interpreted this phenotypic ratio as the result of the segregation of alleles at two loci. (Although he found alleles at three loci that affect kernel color, the two varieties used in this cross differed at only two of the loci.) He proposed that there were two alleles at each locus: one that produced red pigment and another that produced no pigment. We'll designate the alleles that encoded pigment A^+ and B^+ and the alleles that encoded no pigment A^- and B^- . Nilsson-Ehle recognized that the effects of the genes were additive. Each gene seemed to contribute equally to color; so the overall phenotype could be determined by adding the effects of all the genes, as shown in the following table.

Genotype	Doses of pigment	Phenotype
$A^+A^+ B^+B^+$	4	purple
$A^+A^+ B^+B^- \}$	3	dark red
$A^+A^- B^+B^+ \}$	3	red
$A^-A^- B^+B^- \}$	2	light red
$A^+A^- B^+B^- \}$	1	light red
$A^-A^- B^+B^- \}$	0	white

Notice that the purple and white phenotypes are each encoded by a single genotype, but other phenotypes may result from several different genotypes.



To calculate Probability

Probability

In our example, $x_1 = 160$, $x_2 = 161$, $x_3 = 167$, and so forth.
The mean height (\bar{x}) equals:

$$\bar{x} = \frac{160 + 161 + 167 + 164 + 165}{5} = \frac{817}{5} = 163.4$$

A shorthand way to represent this formula is

$$\bar{x} = \frac{\sum x_i}{n}$$

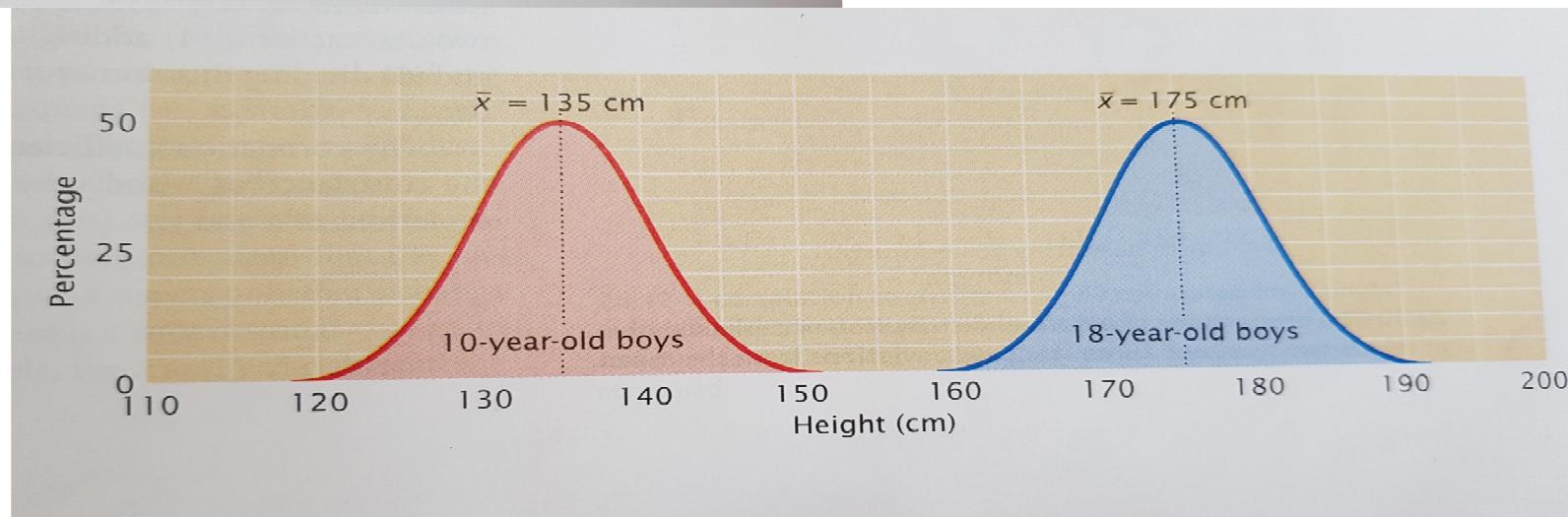
or

$$\bar{x} = \frac{1}{n} \sum x_i$$

where the symbol Σ means “the summation of” and x_i represents individual x values.

The Mean

1. Provides information about the Centre of the distribution
2. If Heights of 10 year old and 18 year old boys are measured and plotted frequency of distribution for each group.
3. Both distributions are normal, but they both would be centred at different heights, which would be indicated by different Mean.



Variation and Standard Deviation

- It indicates variability of a group of measurements or how spread out the distribution is.
- Distributions may have the same mean but different variance.
- Larger the variance, the greater the spread of measurements in a distribution about its mean.

$$s^2 = \frac{\sum(x_i - \bar{x})^2}{n - 1}$$

To calculate the variance, we (1) subtract the mean from each measurement and square the value obtained, (2) add all the squared deviations, and (3) divide this sum by the number of original measurements minus 1.

Another statistic that is closely related to the variance is the **standard deviation** (s), which is defined as the square root of the variance:

$$s = \sqrt{s^2}$$

Correlation

Correlation

The mean and the variance can be used to describe an individual characteristic, but geneticists are frequently interested in more than one characteristic. Often, two or more characteristics vary together. For instance, both the number and the weight of eggs produced by hens are important to the poultry industry. These two characteristics are not independent of each other. There is an inverse relation between egg number and weight: hens that lay more eggs produce smaller eggs. This kind of relation between two characteristics is called a **correlation**. When two characteristics are correlated, a change in one characteristic is likely to be associated with a change in the other.

Correlations between characteristics are measured by a **correlation coefficient** (designated r), which measures the strength of their association. Consider two characteristics, such as human height (x) and arm length (y). To determine how these characteristics are correlated, we first obtain the covariance (cov) of x and y :

$$\text{cov}_{xy} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{n - 1}$$

The covariance is computed by (1) taking an x value for an individual and subtracting it from the mean of x (\bar{x}); (2) taking the y value for the same individual and subtracting it from the mean of y (\bar{y}); (3) multiplying the results of these two subtractions; (4) adding the results for all the xy pairs; and (5) dividing this sum by $n - 1$ (where n equals the number of xy pairs).

The correlation coefficient (r) is obtained by dividing the covariance of x and y by the product of the standard deviations of x and y :

$$r = \frac{\text{cov}_{xy}}{s_x s_y}$$

Correlation Coefficient

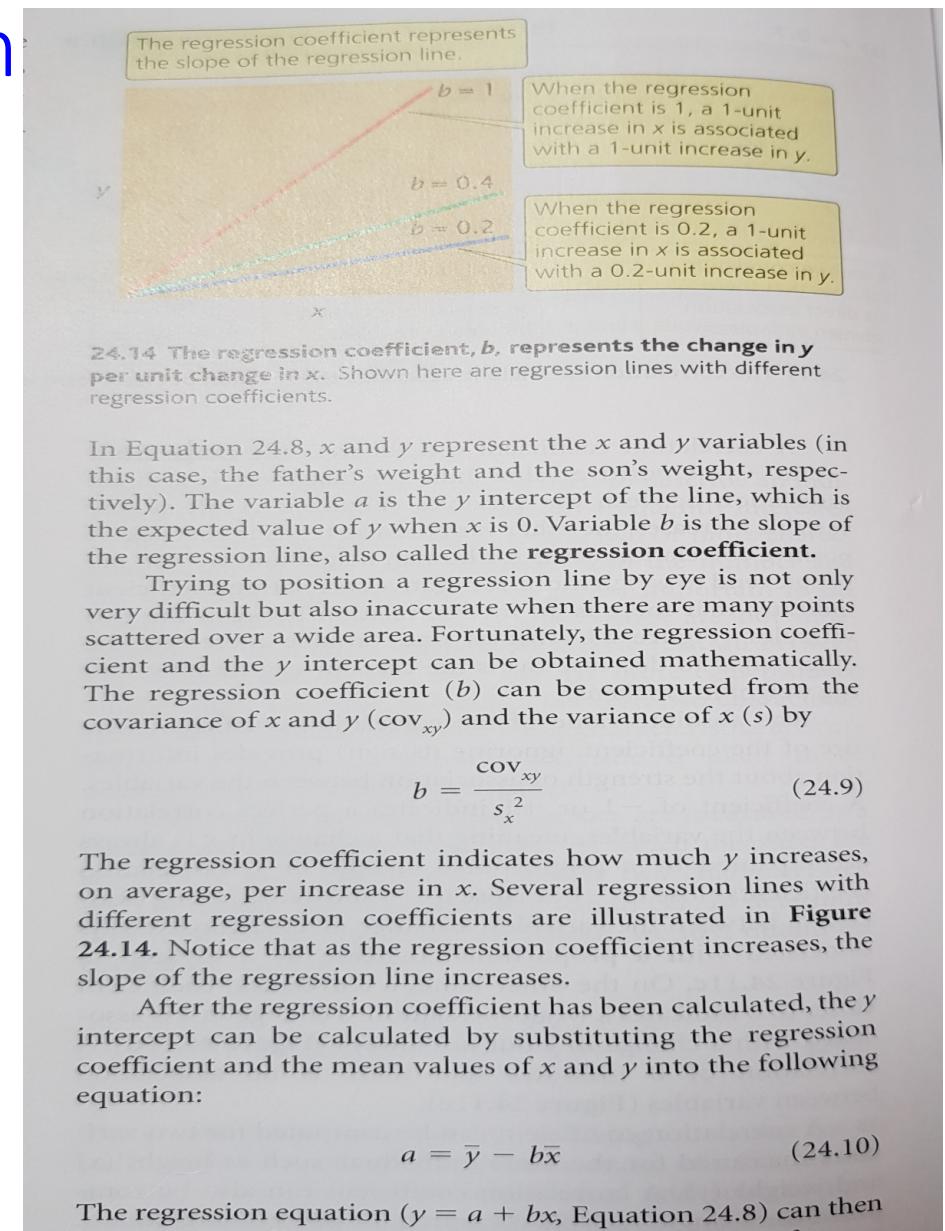
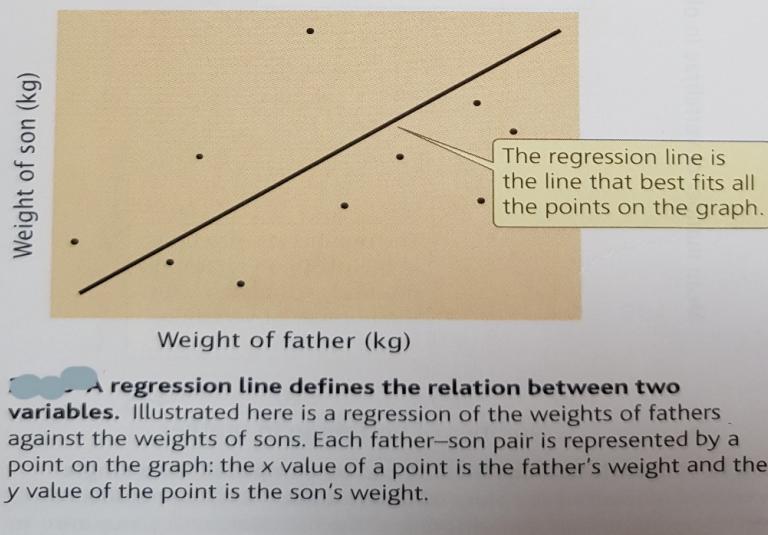
- Correlation coefficient, theoretically range from -1 to +1.
- A positive value indicates direct association between variables, as one variable increases, the other variable also tend to increase. E.g. Human height and weight, taller people tend to weigh more
- Negative value indicates inverse relation between the two variables; when one increases, the other tend to decreases. E.g. Egg number and Egg weight
- The absolute value provides strength of association.
- Correlation Coefficient close to zero (0) indicates a weak correlation, a change in x is associated with a change in y but not always

Regression

A positive correlation exists between the body weight of parents and the body weight of their offspring; this correlation exists in part because genes influence body weight, and parents and children have genes in common. Because of this association between parental and offspring phenotypes, we can predict the weight of an individual on the basis of the weights of its parents. This type of statistical prediction is called **regression**. This technique plays an important role in quantitative genetics because it allows us to predict the characteristics of offspring from a given mating, even without knowledge of the genotypes that encode the characteristic.

Regression can be understood by plotting a series of x and y values. Figure 24.13 illustrates the relation between the weight of a father (x) and the weight of his son (y). Each father–son pair is represented by a point on the graph. The overall relation between these two variables is depicted by the regression line, which is the line that best fits all the points on the graph (deviations of the points from the line are minimized). The regression line defines the relation between the x and y variables and can be represented by

$$y = a + bx \quad (24.8)$$



Question

Body weights of 11 female fishes and the numbers of eggs that they produce are:

Weight (mg)	Eggs (thousands)
x	y
14	61
17	37
24	65
25	69
27	54
33	93
34	87
37	89
40	100
41	90
42	97

What are the correlation coefficient and the regression coefficient for body weight and egg number in these 11 fishes?

Ques help

A Weight (mg)	B	C	D Eggs (thousands)	E	F	G
x	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
14	-16.36	267.65	61	-15.55	241.80	254.40
17	-13.36	178.49	37	-39.55	1564.20	528.39
24	-6.36	40.45	65	-11.55	133.40	73.46
25	-5.36	28.73	69	-7.55	57.00	40.47
27	-3.36	11.29	54	-22.55	508.50	75.77
33	2.64	6.97	93	16.45	270.60	43.43
34	3.64	13.25	87	10.45	109.20	38.04
37	6.64	44.09	89	12.45	155.00	82.67
40	9.64	92.93	100	23.45	549.90	226.06
41	10.64	113.21	90	13.45	180.90	143.11
42	11.64	135.49	97	20.45	418.20	238.04
$\sum x_i = 334$		$\sum (x - \bar{x})^2 = 932.55$	$\sum y_i = 842$		$\sum (y - \bar{y})^2 = 4188.70$	$\sum (x_i - \bar{x})(y_i - \bar{y}) = 1743$