Rules of Probability in Inheritance of Traits/ Genetic Analysis

Rules of Probability in Inheritance of Traits/ Genetic Analyses

Two different approaches of genetic analysis are normally conducted by geneticists frequently:

A) Predicting the genotypes of the parents from the phenotypic class ratio in progeny.

This is sometimes referred as Forward Genetics

B) Predicting phenotypic class ratio in progeny from parents of known genotypes.

This is sometimes known as Reverse Genetics

Rules of Probability in Inheritance of Traits/ Genetic Analysis

Contd..

A. Predicting the genotypes of the parents from the phenotypic class ratio in progeny

A) 1. Use of test cross:

We have seen before how **Mendel's test cross** can be used to decipher the genotype of an individual showing dominant phenotype.

Unknown genotype (*A*_) with phenotype dominant ***A*** allele **X** homozygous recessive (*aa*) **Tester**

If the progeny phenotypic ratio is 1:1 (50% dominant and 50% recessive), then the genotype was ***Aa***

If all individuals of progeny show dominant phenotype, then the genotype was ***AA***

The **tester being fully recessive meaning homozygous recessive with respect to the trait** (or gene) under consideration, **its gametes carry the recessive allele and should not contribute to** the phenotype of the progeny in this generation.

Rules of Probability in Inheritance of Traits/ Genetic Analysis

Contd..

A. Predicting the genotypes of the parents from the phenotypic class ratio in progeny

A) 2. In the absence of a tester, by selfing:

If the unknown genotype, after selfing produces a progeny of phenotypic ratio 3:1, then it is heterozygous for the alleles of that gene.

If the unknown genotype, after selfing produces a progeny of all similar phenotypes, then it is homozygous for the alleles of that gene, i.e. pure-line or true-breeding type.

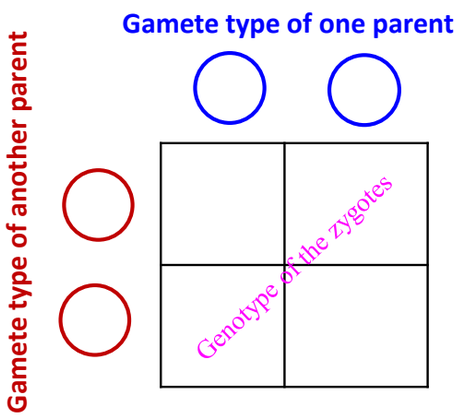
Rules of Probability in Inheritance of Traits/ Genetic Analysis Contd..

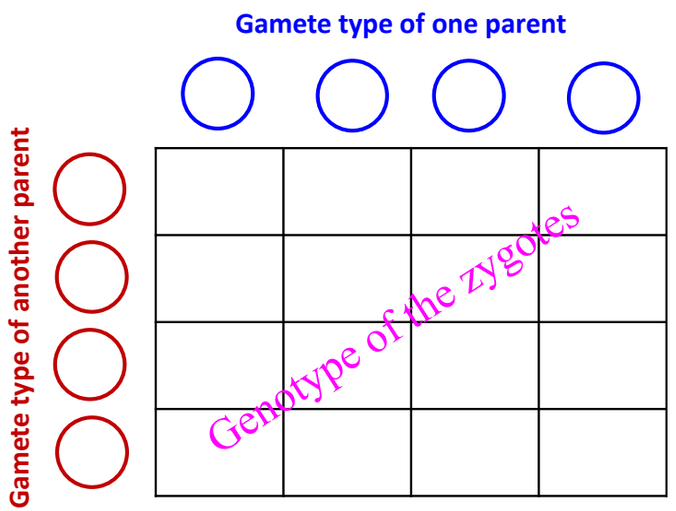
B) Predicting progeny phenotypic class ratio from the cross of two known genotypes is important

- in many breeding programs of plants and animals to develop improved cultivars or stocks with desirable traits.
- in predicting and explaining the inheritance pattern of a disease causing trait in a family or population.

We have used **Punnett square** earlier to explain the Mendel's observation on phenotypic and genotypic ratio for **monohybrid** and **dihybrid crosses**

Depending upon the dominant and recessive alleles in the zygote, we predict the phenotypes of these individuals

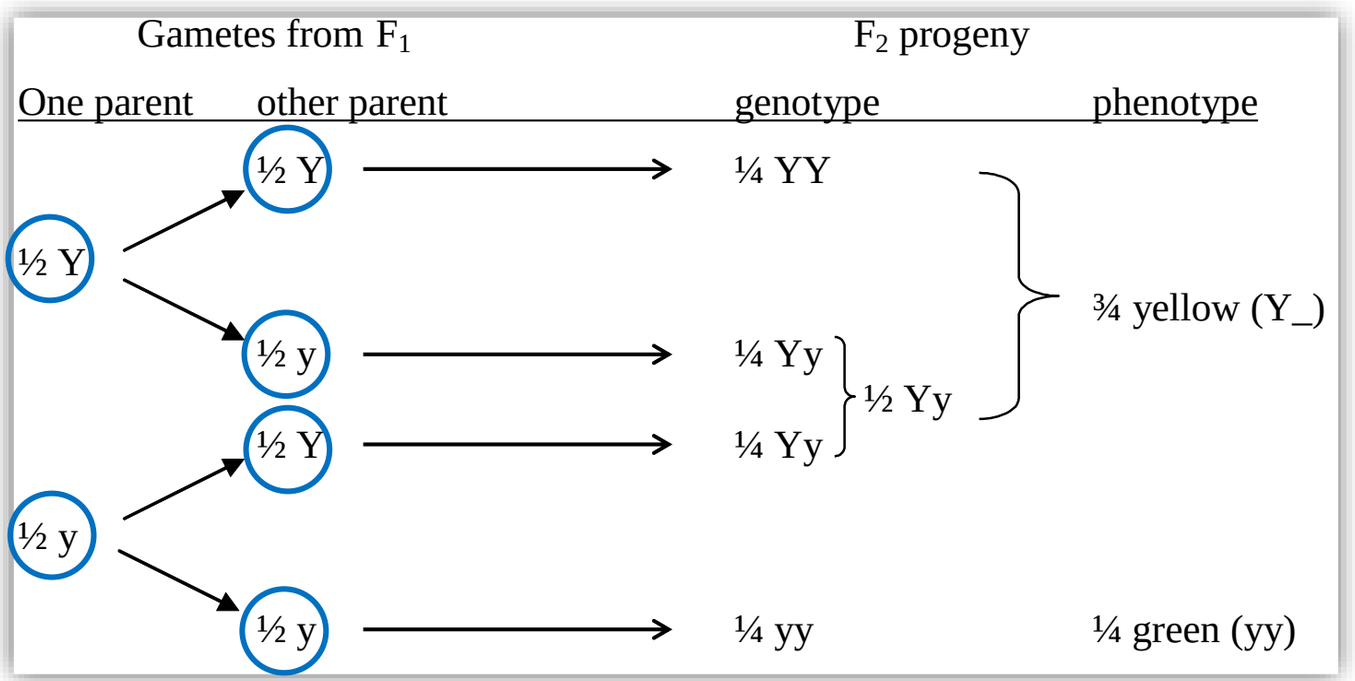




This may be convenient for predicting the inheritance pattern of one or two traits (or genes), but it is **very difficult to draw graphic squares or grids** when several pairs of alleles are involved for more than two genes at a time.

Rules of Probability in Inheritance of Traits/ Genetic Analysis Contd..

An alternate to Punnett square is to use **branch-line diagram**.
For monohybrid selfing (***Yy*** X ***Yy***), it looks like this:



- However, this tree for branch diagram becomes quite **difficult to draw for two or more pairs of alleles controlling more than one trait**.
- Therefore, to simplify both Punnett square and branch-line diagram, **statistical rules of probability** are used to predict the genotypes and phenotypes of the progeny in genetic analysis of the inheritance study.

We need to keep in mind that the Punnett square and branch diagram are actually based upon these probability rules, which are also reflected in genotypic and phenotypic ratios observed in Mendel's experiments.

Rules of probability or laws of chance

Used to **calculate the expected frequencies of types/classes of**

- gametes,
- genotypes or
- phenotypes resulting from a cross of known genotypes.

Two rules are important to understand the inheritance of traits or genes.

1. **Product rule of probability/ Rule of Multiplication**
2. **Sum rule of probability/ Rule of addition**

Rules of probability or laws of chance

Contd..

1. Product rule of probability/ Rule of Multiplication:

It states that the probability of **two independent events both occurring together or simultaneously** is **the product** of their individual probabilities.

Examples: **What is the probability (p) of obtaining a pair of 4's from rolling two dice?**

p (of a 4 in one die) = 1/6

Since, the outcome depends on both events occurring together but independently, using the product rule

p (of two 4's in both dice) = 1/6 x 1/6 = **1/36**

Question: **What is the probability that both children of a couple with two children will be girl?**

Answer: (½ x ½) = **¼**

Rules of probability or laws of chance

Contd..

2. Sum rule of probability/ Rule of addition:

It states that the probability of **either of two mutually exclusive events** occurring is **the sum** of their individual probabilities.

Example: **What is the probability (p) of obtaining either a pair of 4's or a pair of 5's from rolling two dice?**

p (of a pair of 4's in two dice) = 1/36 (as calculated earlier)

So, p (of a pair of 5's in two dice) = 1/36

Since, the outcomes are mutually exclusive; the sum rule is used-

p (of a pair of 4's or 5's in both dice) = 1/36 + 1/36 = 1/18

Question: **What will be the probability of having either both boys or both girls of a couple with two children?**

Answer: (¼ + ¼) = ½

Rules of probability or laws of chance

Contd..

For genetic inheritance study,

- If the foci are dependent on the outcome of **both genes *A* and *B*** (i.e., both of the two genes); the **product rule is applicable**, because inheritance of *A* gene occurs independent of *B* gene.

- If the focus is dependent on the outcome of **either *A* or *a*** (i.e., either of the two alleles of a particular gene), **the sum rule is applicable**, because if "A" is inherited then "a" cannot; as gametes carry only one allele of the two alleles.

- **Mendel's laws of inheritance in conjunction with the statistical rules of probability** provide geneticists the tools for predicting and interpreting the results of genetic crosses.

Rules of probability or laws of chance

Contd..

Number of types/classes of gametes, genotypes and phenotypes obtained from selfing of heterozygotes or F_1 hybrid (if all genes show true dominance-recessive relationship for the inheritance pattern):

No. of gene (each with two segregating alleles)	No. of types of gametes	No. of genotypic classes	No. of phenotypic classes
1	2	3	**2**
2	4	9	**4**
3	8	27	**8**
4	16	81	**16**
n	**2^n**	**3^n**	**2^n**

Practicing a few problems

Problem 1: How many distinct genotypes and phenotypes will a testcross produce for a tetrahybrid *Aa Bb Cc Dd*, i.e. a cross between *Aa Bb Cc Dd* X *aa bb cc dd* ?

Answer: We know from monohybrid testcross- 2 distinct genotypes and 2 distinct phenotypes produced in the progeny

Considering **laws of segregation of alleles and independent assortment of all the four genes** (each gene having two alleles), **the product rule of probability is used to calculate the genotypic and phenotypic classes** of the progeny from tertrahybrid (four genes) testcross-

2 x 2 x 2 x 2 = 2^4 = 16 genotypes and 16 phenotypes, each with a proportion of 1/16

Practicing a few problems

Contd..

Problem 2: How many distinct genotypes and phenotypes will be produced by selfing of the same tetrahybrid *Aa Bb Cc Dd* ?

Answer: We know from F_1 monohybrid (Yy) selfing (i.e., ***Yy*** X ***Yy***), we can get progeny of **3 genotypic classes: *YY*, *Yy*** and ***yy*** **[Homozygous for dominant allele, Heterozygous and Homozygous for recessive allele]** **and 2 phenotypic classes: yellow** and **green** **[Dominant allele and Recessive allele]**

Considering **laws of segregation of alleles and independent assortment of all the four genes** (each gene having two alleles), **the product rule of probability is used to calculate the genotypic and phenotypic classes** of the progeny from tertrahybrid (four genes) selfing-

$3 \times 3 \times 3 \times 3 = 3^4 =$ 81 genotypes

$2 \times 2 \times 2 \times 2 = 2^4 =$ 16 phenotypes

Practicing a few problems

Contd..

Problem 3: What proportion of progeny will be of a specific genotype say, *aa bb cc dd ee* from the cross of two plants with genotypes *Aa bb Cc Dd Ee* X *Aa Bb Cc dd Ee* ?

Answer: Considering individually **five different genes each with two segregating alleles** which can assort independently-

From ***Aa*** X ***Aa*** ¼ of the progeny will be genotype ***aa***
From ***bb*** X ***Bb*** (2/4 =) ½ of the progeny will be genotype ***bb***
From ***Cc*** X ***Cc*** ¼ of the progeny will be genotype ***cc***
From ***Dd*** X ***dd*** (2/4 =) ½ of the progeny will be genotype ***dd***
From ***Ee*** X ***Ee*** ¼ of the progeny will be genotype ***ee***

Thus, the overall probability or expected frequency of obtaining the desired **genotype *aa bb cc dd ee*** (considering all the five genes follow independent assortment or event) will be
¼ X ½ X ¼ X ½ X ¼ = 1/256

Practicing a few problems

Similarly, the rules of probability can be used to calculate the expected frequencies of phenotypic classes or gamete types.

An example of predicting the expected frequency of phenotypic classes from known genotype

Practicing a few problems

Problem 4: A crossing was made using a plant having purple flowers and yellow, round seeds (heterozygous for all three traits/genes) with a plant having purple flowers and green, wrinkled seeds (heterozygous for flower color but homozygous for other two traits).

What proportion of progeny from this cross is predicted to exhibit the *recessive phenotypes for at least two* of these three traits?

[This means, we need to find out the correct genotypes **which are homozygous recessive for at least two genes**]

Practicing a few problems

Problem 4

Answer: Firstly, determine the genotype of the two parents:

PpYyRr X ***Ppyyrr***

(heterozygous for all three traits/genes) X **(heterozygous for flower color but homozygous for other two traits)**

Now, list all the genotypes that will show the ***recessive phenotypes*** ***(so, genotypes should be homozygous recessive)*** ***for at least two*** **traits** from the above cross:

ppyyRr, ppYyrr, Ppyyrr, PPyyrr, ppyyrr

[This last one shows all three recessive traits. Since the condition is *at least two recessive* traits, we need to include this genotype also]

Practicing a few problems

Problem 4

Answer: Now, apply the **product rule and sum rule of probability**, considering the cross for one gene (two alleles) at a time.

[From a cross between ***Aa*** X ***Aa*** ; the expected proportion of **homozygous dominant (*AA*) is** ¼ , **heterozygous (*Aa*) is 2/4** , and **homozygous recessive (*aa*) is** ¼ .

On the other hand, from a cross between ***Aa*** X ***aa*** ; the expected proportion of **heterozygous (*Aa*) is (2/4 =)** ½, and **homozygous recessive (*aa*) is (2/4 =)** ½]

[The two parents are: *PpYyRr* X *Ppyyrr*] Now, progeny genotypes of homozygous recessive phenotypes for <u>at least two</u> traits:-

ppyyRr	¼ (pp) X ½ (yy) X ½ (Rr) = 1/16
ppYyrr	¼ (pp) X ½ (Yy) X ½ (rr) = 1/16
Ppyyrr	2/4 (Pp) X ½ (yy) X ½ (rr) = 2/16
PPyyrr	¼ (PP) X ½ (yy) X ½ (rr) = 1/16
ppyyrr	¼ (pp) X ½ (yy) X ½ (rr) = 1/16

Thus, total proportion of progeny having at least two recessive traits = (summing up above values) 6/16 = 3/8

Up to this slide, Lecture # 7 delivered on 05.09.2023

Methodology of Science: Hypothetico-deductive method of investigation

1) Characterization/ observation of phenomenon and collection of data

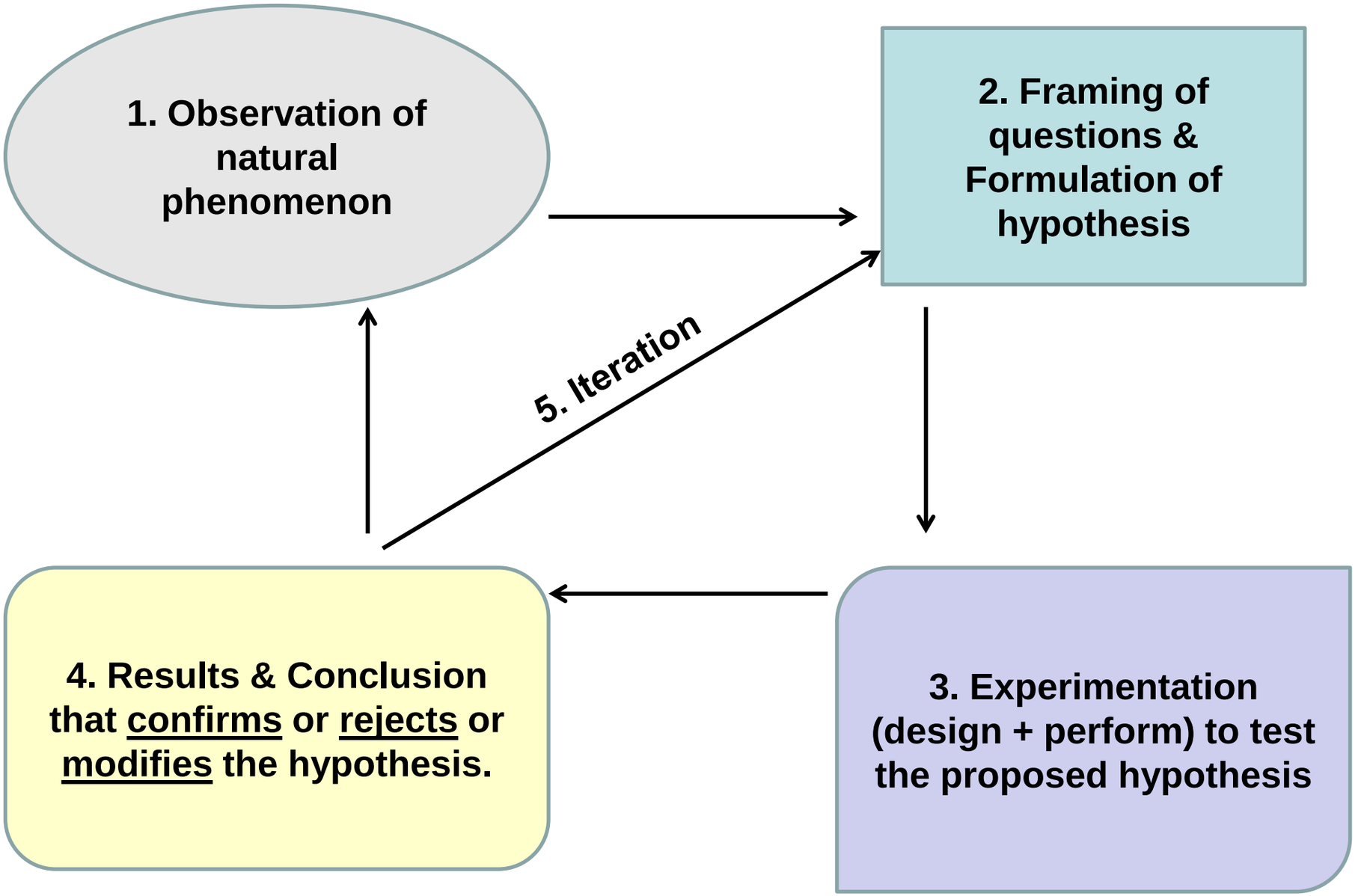
2) Asking questions, formulation of hypothesis, making prediction or designing model

3) Experimentation to verify the hypothesis/ prediction/ model

4) Establish the theory/law (if hypothesis is accepted), or

5) Repeat the cycle with new observation, collection of data, formulation of new hypothesis/making new prediction/model and testing of new hypothesis/prediction/model.

Methodology of Science: Hypothetico-deductive method of investigation Contd..



Ask students about: Mode, Median, Mean, Variance, Standard Deviation, Standard Error

Explain when to use:

1. **Chi-square (χ^2) test**: Quantify & compare the differences **between observed frequencies and expected frequencies.** Thus, it calculates/determines the deviation

2. ***t*-test**: Quantify & compare the differences **between two groups** on some variable of interest (two sets of data).

- **Unpaired *t*-test** (also known as **Student *t*-test**): used for two independent subjects, e.g., diabetic and non-diabetic people
- **Paired t-test**: used for same subject at two different times (paired observations), e.g., before and after treatment using drug/medicine to the diabetic people

3. **ANOVA**: Quantify & compare the differences **among three or more groups** on some variable of interest (three or more sets of data)

Chi-square (χ^2) test, *t*-test & ANOVA

Contd..

- In all the above 3 types of analyses, changes in measurements are considered statistically significant at $p \leq 0.05$ (95% confidence level). However, the χ^2 deviation (difference between observed and expected values) is insignificant.
- One important consideration in these statistical analyses: Final outcome (acceptance or rejection) depends on **Null Hypothesis** (≡ counter claim). It could be framed in two ways:

1) There is **no real difference** between the observed & expected results (e.g., during Mendelian genetic investigation & analysis)

2) There is **indeed difference** between the observed & expected results (e.g., in genetic transformation experiment & analysis)

About *p*-value setting

During the test of Null Hypothesis, a ***p***-value determines the significance of our results.

1) a ***p***-value is number **between 0.0 to 1.0**

2) A **small *p*-value (typically ≤ 0.05)** indicates strong evidence against the Null Hypothesis. Hence, it is to be rejected in favour of an alternative hypothesis

3) In other words, a **high *p*-value (≥ 0.05)** indicates strong evidence in favour of the Null Hypothesis. Hence, the high probability that the Null Hypothesis is true.

4) Hence, we fix the ***p*-value level at 0.05** in a χ^2 table values

5) **If the calculated value is less than (<) *p*(0.05) table value, we say insignificant difference or deviation between observed and expected values**

Statistical analysis of the Mendelian monohybrid and dihybrid ratios by chi-square (χ^2) test

Objectives:

- To understand the importance of statistical test i.e., **chi-square (χ^2) test** in genetic analysis
- To calculate χ^2 value to determine whether a given set of experimental data fit to an expected ratio predicted by a hypothesis (here Mendelian monohybrid and dihybrid ratios)
- To interpret the calculated χ^2 value with a given degree of freedom (df) using the tabulated χ^2 value at certain probability level

What will you learn today?

Calculate the χ^2 value and interpret the results obtained from the data provided to you.

Statistical analysis of the Mendelian monohybrid and dihybrid ratios by chi-square (χ^2) test

Contd..

In genetic crosses, researchers record data that are quantitative i.e., varied number of different classes of phenotypes. It is necessary and important **to check or verify whether the observed results are close to the expected one predicted by the hypothesis being tested**.

- To avoid the problems due to bias or errors; a statistical test, known as **chi-square (χ^2) test** is used to quantify the deviations (differences between observed and expected) if the hypothesis is acceptable or not.
- The chi-square (χ^2) value is the measurement of various deviations expected by random chance (not by errors or bias).
- Thus, chi-square (χ^2) test is a ***goodness-of-fit*** test for verifying the **null hypothesis,** which states that there is no real difference between the expected and observed results.
- If the **chi-square (χ^2) value is insignificant** (at 0.05 probability value i.e., 95% confidence level) then the deviations are due to chance and the **null hypothesis is accepted** as the observed data can be explained by formulated hypothesis.
- On the other hand, **if the chi-square (χ^2) value is significant** then the deviations are not due to chance but because of some other factors and **then the null hypothesis is rejected**, and we need to formulate a new null hypothesis.

Four important considerations/concepts for performing the chi-square (χ^2) test:

1) Framing or formulating a suitable hypothesis called **null hypothesis** (≡ counter claim) which states that there is no real difference between the observed data and the predicted data under the circumstances of equal segregation of alleles and equal viability of the progeny.

2) Sample size (number in each class) should be appropriate i.e., progeny population should be as large as possible.

3) Comparison or calculation of percentage or proportion or ratio of different classes **will not permit** us to determine whether or not the observed results are significantly different from the predicted values.

4) The **absolute numbers** of the observed classes are important and **used for calculating** chi-square (χ^2) value, because they reflect the actual data and authentically tell us if the proposed hypothesis is correct or not.

Testing Mendelian monohybrid and dihybrid ratios by chi-square (χ^2) test:

- Any of the Mendelian ratios obtained from **monohybrid cross** i.e., 3:1 (1 df), 1:1 (1 df), **or dihybrid cross** i.e., 9:3:3:1 (3 df) and 1:1:1:1 (3 df) can be tested using chi-square (χ^2) test.

- We will have more applications of chi-square (χ^2) test for genetic analysis in linkage and recombination study.

The general formula for calculating χ^2 value:

If '**O**' be the observed number of a phenotypic class and '**E**' be the expected number of the same class (based on the ratio predicted by the hypothesis) then,

$$\chi^2 = \Sigma \{(O-E)^2/E\}$$

for all classes with (n-1) degrees of freedom (df), where n = number of classes

Sometimes (O-E) is denoted as deviation or '**d**' and the formula is modified accordingly. The df is the number of independently varying parameters in the experiment.

Testing Mendelian monohybrid and dihybrid ratios by chi-square (χ^2) test:

Contd..

- The ***calculated value*** of χ^2 at the required **df** is then compared with the ***table value*** **at 0.05 probability (*p*) level**.
- ***If the calculated value is less than (<) the table value at 0.05 probability (p) level, then it is said to have insignificant deviation i.e., good fit to the predicted ratio and the null hypothesis is accepted*.**

Need to make a Table:

Class	Observed Number (O)	Expected Number (E)	$(O-E)^2$	$(O-E)^2/E$
1				
2 etc.				

$$\chi^2 = \Sigma \{(O-E)^2/E\}$$

Table: Chi-Square Probabilities (shortened)

df	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
1	---	---	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188

Practicing a few problems

Q1. Upon selfing the F_1 monohybrid tall plants segregated into two phenotypic classes- tall and dwarf comprising of 747 and 253 plants, respectively in the F_2 generation.

(i) Does it follow Mendelian segregation ratio?
(ii) Suggest genotype of parents, F_1 and F_2 plants.
(iii) What is the expected number of homozygous and heterozygous tall plants?

[Answer clue: Total 1000; $\chi^2 = 0.048$]

Practicing a few problems

Q2. In F_2 population of a dihybrid F_1 plant, the following phenotypic classes were observed:

Tall plant with round seeds	**206**
Tall plant with wrinkled seeds	**65**
Dwarf plant with round seeds	**83**
Dwarf plant with wrinkled seeds	**30**

(i) Does it follow the Mendelian law of independent assortment?

(ii) Suggest genotype of parents, F_1 and F_2 plants.

[Answer clue: Total 384; $\chi^2 = 4.32$]

Practice solving questions from the 2-3 Genetics books

Genetics & Genetic Engg. Assignment for 3rd Year BT students (Autumn-2023)

Sl. No.	Topic	Group	Roll No.	Grade
1	Cytoplasmic Inheritance and Male Sterility			
2	DNA fingerprinting			
3	Epigenetics			
4	Genetic Imprinting			
5	Transposable Elements			
6	Genetic Diseases and Gene Therapy			
7	Restriction enzymes empower both bacteria and human			
8	The power of PCR beyond amplification			
9	Revolution in genetic engineering by CRISPR-Cas system			
10	Human insulin as therapeutics: a gift of genetic engineering			

Group formation & allotment of topics on 05/09/2023

Date of presentation **on 30/10/2023 from 3:00 PM to 5:00 PM**

Up to this slide, Lecture # 8 delivered on 05.09.2023