BULLETIN OF MATHEMATICS AND STATISTICS RESEARCH
A Peer Reviewed International Journal,
Contents available on www.bomsr.com
Vol.4. S1.2016; ISSN: 2348-0580
Email:editorbomsr@gmail.com
RADMAS– 2016

APPLICATION OF PRINCIPLES OF PROBABILITY IN MENDELIAN INHERITANCE

Dr. K. MANIKYA KUMARI
Head, Department of Botany
St. Joseph's College for Women (A)Visakhapatnam-530 004, Andhra Pradesh, India
Email: mk.kambham@gmail.com

ABSTRACT

Probabilities are mathematical measures of likelihood. The empirical probability of an event is calculated by dividing the number of times the event occurs by the total number of opportunities for the event to occur. Empirical probabilities come from observations such as those of Mendel. He demonstrated that the different characteristics of pea-plant he studied were transmitted independently of one another from parent to offspring. The probability of two independent events occurring together can be calculated by multiplying the individual probabilities of each event occurring alone. Considering two mutually-exclusive outcomes that can result from more than one pathway, the probability of the occurrence of one event or the other, of two mutually-exclusive events, is the sum of their individual probabilities. In a genetic cross between a dominant and recessive traits, the probability of the dominant trait being expressed is dependent upon its frequency. Many quantitative traits have approximately Normal distributions, because they are the net result of a number of small effects. To use probability laws in practice, it is necessary to work with large sample sizes because small sample sizes are prone to deviations caused by chance. This helps to predict the genetic traits of the offspring accurately even before fertilization, when the parental traits were known. Predicting the expected phenotypes and genotypes of heritable human genetic disorders or diseases helps to eliminate such genes even before their expression through gene therapy or gene editing techniques. This paper discusses the application of rules of probability and Chi- square test and to predict the phenotype and genotype expression of genetic traits and the amount of genetic variation in a population using some heritable traits of plants, animals and humans as examples.

Keywords: Genetic traits, Phenotype, Genotype, Offspring, Probability, Chi – square test

1. Introduction

Probabilities are mathematical measures of likelihood. The empirical probability of an event is calculated by dividing the number of times the event occurs by the total number of opportunities for the event to occur. Empirical probabilities come from observations such as those of Mendel. An example of a genetic event is a round seed produced by a pea plant. Mendel demonstrated that the probability of the event "round seed" was guaranteed to occur in the F_1 offspring of true-breeding parents, one of which has round seeds and one of which has wrinkled seeds. When the F_1 plants were subsequently self-crossed, the probability of any given F_2 offspring having round seeds was now three out of four. In other words, in a large population of F_2 offspring chosen at random, 75 percent were expected to have round seeds, whereas 25 percent were expected to have wrinkled seeds. Using large numbers of crosses, Mendel was able to calculate probabilities and use these to predict the outcomes of other crosses.

BULLETIN OF MATHEMATICS AND STATISTICS RESEARCH
A Peer Reviewed International Journal,
Contents available on www.bomsr.com
Vol.4. S1.2016; ISSN: 2348-0580
Email:editorbomsr@gmail.com
RADMAS– 2016

2. Rules of Probability

2. 1. The Product Rule: Mendel demonstrated that the pea-plant characteristics he studied were transmitted as discrete units from parent to offspring. Mendel also determined that different characteristics were transmitted independently of one another and could be considered in separate probability analyses. For instance, performing a cross between a plant with green, wrinkled seeds and a plant with yellow, round seeds produced offspring that had a 3:1 ratio of green : yellow seeds and a 3:1 ratio of round : wrinkled seeds. The characteristics of color and texture did not influence each other.

The product rule of probability can be applied to this phenomenon of the independent transmission of characteristics. It states that the probability of two independent events occurring together can be calculated by multiplying the individual probabilities of each event occurring alone. Imagine that you are rolling a six-sided die (D) and flipping a penny (P) at the same time. The die may roll any number from 1–6 ($D_{\#}$), whereas the penny may turn up heads (P_H) or tails (P_T). The outcome of rolling the die has no effect on the outcome of flipping the penny and vice versa. There are 12 possible outcomes, and each is expected to occur with equal probability: D_1P_H, D_1P_T, D_2P_H, D_2P_T, D_3P_H, D_3P_T, D_4P_H, D_4P_T, D_5P_H, D_5P_T, D_6P_H, $D_6\,P_T$.

When we flip two coins, they have the possibility of landing heads/heads, heads/tails, tails/heads, or tails/tails. Each of these is equally possible to happen, so each represents one-fourth of all the possible occurrences. These values are also the probabilities of the events occurring. The probability of flipping two coins landing:

heads/heads 1/4, tails/tails 1/4, and heads with tails in any order 1/4 1/4 1/4 1/2

The genetic probability of traits is similar to the probability of flipping a coin. In human genetics, the egg and sperm each supply the baby one of two pieces of genetic information per trait from its parents. These pieces of genetic information are called alleles. Alleles are either dominant or recessive. Dominant alleles are indicated by capital letters, whereas recessive alleles are indicated by lower case letters. Let's look at the probability of a baby inheriting either ridges or no ridges on its fingernails. We'll use the letter "R" to represent the dominant allele, and "r" to represent the recessive allele. Suppose the genetic make-up of both parents is Rr. Each allele has an equal chance of being passed to the child, so the probability for 1/2, and the probability for 1/2(just as in flipping a coin). probability of genetics works as below.

$$probabilit_{RR} = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{4}$$

$$probabilit_{} = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) + \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{4} + \frac{1}{4} = \frac{2}{4} = \frac{1}{2}$$

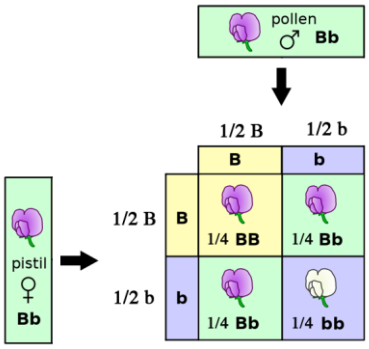
$$probabilit_{rr} = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{4}$$

*Note: The sum of all the probabilities for allele combinations (as above) is one.

Of the 12 possible outcomes, the die has a 2/12 (or 1/6) probability of rolling a two, and the penny has a 6/12 (or 1/2) probability of coming up heads. The probability that you will obtain the combined outcome 2 and heads is: (D_2) x (P_H) = (1/6) x (1/2) or 1/12. The word "and" is a signal to apply the product rule. Consider how the product rule is applied to a dihybrid : the probability of having both

BULLETIN OF MATHEMATICS AND STATISTICS RESEARCH
A Peer Reviewed International Journal,
Contents available on www.bomsr.com
Vol.4. S1.2016; ISSN: 2348-0580
Email:editorbomsr@gmail.com
RADMAS– 2016

dominant traits in the F_2 progeny is the product of the probabilities of having the dominant trait for each characteristic.



Example: 1. Role of probability in segregation of alleles and fertilization

In a genetic cross, the probability of the dominant trait being expressed is dependent upon its frequency. In this case, both parents possessed a dominant and a recessive gene for the trait of flower color. The dominant trait is expressed in 3/4 of the offspring and the recessive trait is expressed in 1/4.

2.2. The Sum Rule: The sum rule is applied when considering two mutually-exclusive outcomes that can result from more than one pathway. It states that the probability of the occurrence of one event or the other, of two mutually-exclusive events, is the sum of their individual probabilities. The word "or" indicates that you should apply the sum rule. Let's imagine you are flipping a penny (P) and a quarter (Q). What is the probability of one coin coming up heads and one coming up tails? This can be achieved by two cases: the penny is heads (P_H) and the quarter is tails (Q_T), or the quarter is heads (Q_H) and the penny is tails (P_T). Either case fulfills the outcome.

We calculate the probability of obtaining one head and one tail as

$[(P_H) \times (Q_T)] + [(Q_H) \times (P_T)] = [(1/2) \times (1/2)] + [(1/2) \times (1/2)] = 1/2.$

The sum rule can be applied to show the probability of having just one dominant trait in the F_2 generation of a dihybrid cross.

To use probability laws in practice, it is necessary to work with large sample sizes because small sample sizes are prone to deviations caused by chance. The large quantities of pea plants that Mendel examined allowed him to calculate the probabilities of the traits appearing in his F_2 generation. This discovery meant that when parental traits were known, the offspring's traits could be predicted accurately even before fertilization.

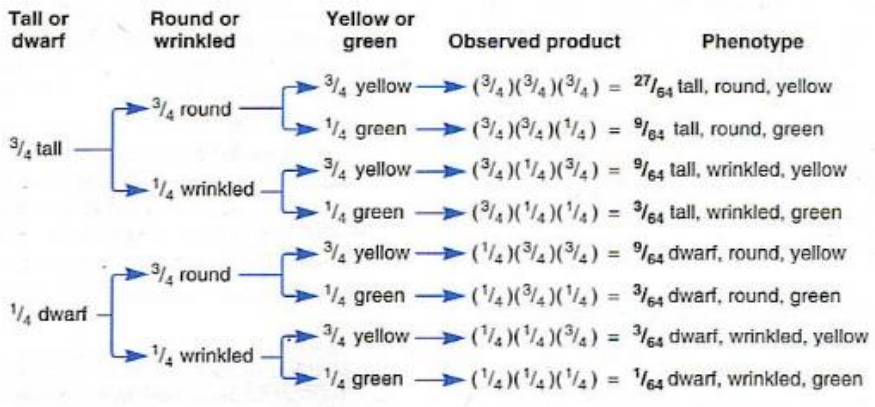
Example: 2. A Pea plant is heterozygous for three genes (Tt Rr Yy), where T=tall, t=dwarf, R=round seeds, r=wrinkled seeds, Y=yellow seeds and y=green seeds. If this plant is self-fertilized, what are the predicted phenotypes of the offspring and what fraction of the offspring will occur in each category? What fraction of progeny will be tall, round and green?

Answer: A cross is made between Tt Rr Yy x Tt Rr Yy. The predicted phenotypes of the offspring and fraction of the offspring will occur in each category as follows.

The fraction of tall, round and green progeny is 9/64. The analysis of the cross was shown using the fork line method.

BULLETIN OF MATHEMATICS AND STATISTICS RESEARCH
A Peer Reviewed International Journal,
Contents available on www.bomsr.com
Vol.4. S1.2016; ISSN: 2348-0580
Email:editorbomsr@gmail.com
RADMAS– 2016

Alternative = Forked Line Method



2.3. Binomial Expansion Equation

Represents all of the possibilities for a given set of unordered events

$$P = \frac{n!}{x!\,(n-x)!}\, p^x\, q^{n-x}$$

Where,

P = probability that the unordered number of events will occur

n = total number of events

x = number of events in one category

p = individual probability of x

q = individual probability of the other category

Note:

p + q = 1

The symbol ! denotes a factorial

n! is the product of all integers from n down to 1

– 4! = 4 X 3 X 2 X 1 = 24

– An exception is 0! = 1

2.4. Probability in genetics : For Independent events, multiply probabilities (1/2 x 1/2 =1/4) For Mutually Exclusive events, add probabilities;

(1/8+1/8+....)

the total probability of all possible outcomes must Sum to one.

Eg.1 : Probability of Child Gender:

i) For a couple with 1 child

- Probability of a boy = ½
- Probability of a girl = ½ (= 1 - probability of a boy)

ii) For a couple with 2 children:

- Probability of 2 boys = ½ x ½ = ¼
- Probability of 2 girls = ½ x ½ = ¼

BULLETIN OF MATHEMATICS AND STATISTICS RESEARCH
A Peer Reviewed International Journal,
Contents available on www.bomsr.com
Vol.4. S1.2016; ISSN: 2348-0580
Email:editorbomsr@gmail.com
RADMAS– 2016

Probability of 1 boy and 1 girl = (1 - probability of not having 2 boys or 2 girls) = 1 - ¼ - ¼ = ½

-or-

- First having a boy then a girl (½ x ½ = ¼) + First having a girl then a boy (½ x ½ = ¼) = ½

Eg. 2: **Probability of inheriting a genetical disorder**

a. In genetically inherited disease such as cystic fibrosis If both parents are carriers of the recessive
allele for a disorder, all of their children will face the following odds of inheriting it:
25% chance of having the recessive disorder
50% chance of being a healthy carrier
25% chance of being healthy and not have the recessive allele at all

b. If one parent is a carrier and the other has a recessive disorder, their children will have the following odds of inheriting it:
50% chance of being a healthy carrier
50% chance having the recessive disorder

c. If only one parent has a single copy of a dominant allele for a dominant disorder, their children will have a 50% chance of inheriting the disorder and 50% chance of being entirely normal.

3. The Chi Square Test

It is a statistical method used to determine goodness of fit. Goodness of fit refers to how close the observed data are to those predicted from a hypothesis. Any statistical test that uses the chi square distribution can be called chi square test. It is applicable both for large and small samples-depending on the context.

3. 1. Tests for Different Purposes

There are different types of chi square test each for different purpose. Some of the popular types are outlined below

1. *Chi square test for testing goodness of fit* is used to decide whether there is any difference between the observed (experimental) value and the expected (theoretical) value.
2. *Chi square test for independence of two attributes.*
3. *Chi square test for single variance* is used to test a hypothesis on a specific value of the population variance.

3. 2. Chi-squared test for variance in a normal population

If a sample of size n is taken from a population having a normal distribution, then there is a result which allows a test to be made of whether the variance of the population has a pre-determined value. For example, a manufacturing process might have been in stable condition for a long period, allowing a value for the variance to be determined essentially without error. Suppose that a variant of the process is being tested, giving rise to a small sample of n product items whose variation is to be tested. The test statistic T in this instance could be set to be the sum of squares about the sample mean, divided by the nominal value for the variance (i.e. the value to be tested as holding). Then T has a chi-squared distribution with n − 1 degrees of freedom. For example, if the sample size is 21, the acceptance region for T with a significance level of 5% is between 9.59 and 34.17.

BULLETIN OF MATHEMATICS AND STATISTICS RESEARCH
A Peer Reviewed International Journal,
Contents available on www.bomsr.com
RADMAS– 2016
Vol.4. S1.2016; ISSN: 2348-0580
Email:editorbomsr@gmail.com

Note:

– The chi square test does not prove that a hypothesis is correct

• It evaluates whether or not the data and the hypothesis have a good fit.

The Chi Square Test general formula is

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

Where,

O = observed data in each category

E = observed data in each category based on the experimenter's hypothesis

Σ= Sum of the calculations for each category

Example: Consider the following example in Drosophila melanogaster

• Gene affecting wing shape
C^+= Normal wing
c= Curved wing

• Gene affecting body color
e+= Normal (gray)
e = ebony

Note: The wild-type allele is designated with a + sign
Recessive mutant alleles are designated with lower case letters

Answer: A cross is made between two true-breeding flies ($C^+C^+e^+e^+$ and ccee). The flies of the F1 generation are then allowed to mate with each other to produce an F2generation

F1 generation : All offspring have straight wings and gray bodies

F2 generation:

- 193 straight wings, gray bodies
- 69 straight wings, ebony bodies
- 64 curved wings, gray bodies
- 26 curved wings, ebony bodies
- 352 total flies

Applying the chi square test

Step 1: Propose a hypothesis that allows us to calculate the expected values based on Mendel's laws. The two traits are independently assorting

Step 2: Calculate the expected values of the four phenotypes, based on the hypothesis

Step 3: Apply the chi square formula

$$\chi^2 = \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2} + \frac{(O_3 - E_3)^2}{E_3} + \frac{(O_4 - E_4)^2}{E_4}$$

Step 4: Interpret the chi square value

The calculated chi square value can be used to obtain probabilities, or P values, from a chi square table

These probabilities allow us to determine the likelihood that the observed deviations are due to random chance alone

– Low chi square values indicate a high probability that the observed deviations could be due to random chance alone

BULLETIN OF MATHEMATICS AND STATISTICS RESEARCH
A Peer Reviewed International Journal,
Contents available on www.bomsr.com
RADMAS– 2016
Vol.4. S1.2016; ISSN: 2348-0580
Email:editorbomsr@gmail.com

– High chi square values indicate a low probability that the observed deviations are due to random chance alone

– If the chi square value results in a probability that is less than 0.05 (ie: less than 5%)

The hypothesis is rejected

Before we can use the chi square table, we have to determine the degrees of freedom (df). The df is a measure of the number of categories that are independent of each other

df = n –1

where,

N = total number of categories

In our experiment, there are four phenotypes/categories

Therefore, df = 4 – 1 = 3

Refer to Table below

Chi Square Values and Probability

Degrees of Freedom	P = 0.99	0.95	0.80	0.50	0.20	0.05	0.01
1	0.000157	0.00393	0.0642	0.455	1.642	3.841	6.635
2	0.020	0.103	0.446	1.386	3.219	5.991	9.210
3	0.115	0.352	1.005	2.366	4.642	7.815	11.345
4	0.297	0.711	1.649	3.357	5.989	9.488	13.277
5	0.554	1.145	2.343	4.351	7.289	11.070	15.086
6	0.872	1.635	3.070	5.348	8.558	12.592	16.812
7	1.239	2.167	3.822	6.346	9.803	14.067	18.475
8	1.646	2.733	4.594	7.344	11.030	15.507	20.090
9	2.088	3.325	5.380	8.343	12.242	16.919	21.666
10	2.558	3.940	6.179	9.342	13.442	18.307	23.209
15	5.229	7.261	10.307	14.339	19.311	24.996	30.578
20	8.260	10.851	14.578	19.337	25.038	31.410	37.566
25	11.524	14.611	18.940	24.337	30.675	37.652	44.314
30	14.953	18.493	23.364	29.336	36.250	43.773	50.892

4. Conclusion: Probability is a bio-statistical calculation used to predict expected phenotypes and genotypes of offspring in Mendelian inheritance. Probability gives a clue to detect heritable genetic disorders before they being expressed. Some of such unwanted heritable traits can be avoided by predetermining their expected genotypes. This utilizes various rules of probability and Chi-squared test analysis.

5. References

[1]. "Rules of Probability for Mendelian Inheritance." *Boundless Biology*. Boundless, 26 May. 2016. Retrieved 20 Oct. 2016

[2]. Yates, F (1934). "Contingency table involving small numbers and the χ^2 test". Supplement to the Journal of the Royal Statistical Society **1**(2): 217–235. JSTOR 2983604

[3]. "Chi-squared Statistic". *Practical Cryptography*. Retrieved 18 February 2015.

[4]. "Using Chi Squared to Crack Codes". *IB Maths Resources. British International School Phuket.*

[5]. *Ryabko, B.Ya.; Stognienko, V.S.; Shokin, Yu.I. (2004).* "A new test for randomness and its application to some cryptographic problems" *(PDF). Journal of Statistical Planning and Inference. **123**: 365–376*. Retrieved 18 February 2015.