3. Test of Hardy – Weinberg equilibrium

If data are available for a locus where all the genotypes are recognizable, the observed frequencies of the genotypes can be tested if they closely agree with a population in H – W equilibrium. According to H – W law, the genotype frequencies of the progeny are determined by the gene frequency of their parents. The gene frequency is the same in parents and progeny, so the gene frequency observed in the progeny can be used as if it were a parental gene frequency to calculate the genotype frequencies expected by H – W law.
A statistical test (chi-square) or X2 is used to determine if the observed and expected genotypic numbers are quite close or significantly different i.e. the X2 tests how well, or how badly the observed numbers agree with the expected.

The following numbers of the human M –N blood groups were recorded in a sample from Iceland. MM = 233, MN = 385, NN = 129

i. What are the genotype frequencies expected from H-W law
ii. How well do the observed number agree with the expectation

Solution	**Genotypes**				**Gene frequencies**	
	MM	MN	NN	Total	M	N
Numbers observed	233	385	129	747	0.57	0.43
Number expected	(57)2 x 747		2pq(747)		(.43)2 x 747	
	= 242.36		= 366.26		= 138.38	

$X2 = \frac{(o - e)2}{e}$ 1.96 P = 2.0

* Table x 2 at 1 df = 3.84

To determine if 1.96 is statistically significant, DF has to be known = 3 -1 -1 =1 (ie No. of categories -1; additional one is subtracted because we had to estimate allelic frequency in which case, only q because p = I-q) in order to calculate the expected number of genotypes. Because 1.96 < 3.84 at one df, it is concluded that the observed numbers in the sample are consistent with those expected under H – W principle. The discrepancy (insignificant) could have arisen by chance in the sampling.

Multiple alleles – consequence – most individuals in population are hets.
Many 2 alleles examined by protein electrophoresis have more than 2 alleles. The frequency of an allele can be calculated as the sum of the frequency of its homozygote + half the frequency of each het. That carries the allele. P = NAiAi + ½ NA1A2 (N = sample of N individuals) Ex. The ABO blood group in man are determined by a series of alleles genes which are 3: A, B, O. Since O is recessive to both A & B H – W frequencies have to be assumed for gene frequencies.

Let the frequencies of A, B & o genes = p, q, & r respectively
Then p + q + r = 1

Blood group	**Genotype**	**Expected**	**Observed %**

A	AA + AO	P2 + 2pr	41.72
B	BB + BO	q2 + 2qr	8.56
O	OO	r2	46.68
AB	AB	2pq	3.04

Overall the genotypic frequencies are equal to the square of the following trinomial

$(p + q + r)2 = p2 + 2pq + q2 + 2pr + 2qr + r2$

Frequency of O gene = r2 using H – W eqm.
Sum of the frequencies of B & O groups = q2 + 2qr + r2
= (q + r)2 = (1 – p)2
= p = 1 (B + O) = B & O & frequencies of blood group.
From table above q = 1 – (0.417 + 0.467) = 0.0598

B gene: p = 1 – (0.0856 + 0.467) = 0.2567
O gene: r = 0.0304 = 0.68

H = 2 pq + 2pr + 2qr = 48

Simples approach is to calculate H –W homozygosity and subtract from 1. The remainder is H – W heterozygosity.

An important consequence of having many alleles at a locus is that most individuals in a population will be heterozygotes. To illustrate this,

H – W heterozygosity for the ABO system

H = 2pq + 2pr + 2qr
= 0.48

When there are many alleles (more hets than homozygotes), the simplest approach is to calculate H – W homozygosity and subtract it from one. The remainder is H – W heterozygosity.

Sex Linked genes

A number of human diseases haemophilia colour blindness and muscular dystrophy) are determined by recessive alleles on the x chromosome. The diseases are much more common in males, than in females on which can be explained by H –W principles.

Males need only one copy of the defective allele to express the trait, while females need two when alleles is rare, it is unlikely that homozygous qs will be encountered.

	A(p)	a (q)
A(p)	AA(p)	Aa pq
a(q)	Aapq	aaq2

From the above, the frequency of affected males, Xa Y is q, while the frequency of affected female Xa Xa = q2. If q is low, it is for most disease alleles (say, 0.005), the frequencies of affected males and qs are 0.005 and 0.000025, respectively. Affected males to female ration is q1q2or 1:q if q = 0.005, the ratio = 200:1 (.005)

$(.005)^2$

Nearly 1:4 in some populations are male hemophilics, compared to qs. The relationship between gene and genotype frequency in the homogametic sex is the same as with an autosonal gene, but the treterogametic sex has only 2 genotypes and each individual carries only one gene instead of 2.

	Females			Male	
	AA	Aa	aa	A	a
	xx	xx	xx	xy	xy
Frequency	P	H	Q	R	S

Hence 2/3 of sex linked genes in the population are carried by the homogametic sex and 1/3 by heterogametic sex. Therefore frequency of A among female population.
Frequency A (Pf) = P + ½ H; frequency A among female Pm = R

Frequency A in the entire population P = 2/3 P+ 1/3 P+ + 1/3 Pm

= 2/3 (P+ + ½ H) + 1/3R = 2/3 Pf + 1/3H + 1/3R

= 1/3 (2P + H + R)

If the gene frequencies among males and females are different, the population is not in equilibrium. The gene frequency in the population as a whole does not change, but its distribution between the 2 sexes oscillates as the population approaches eqm. The reason

Is that males get their sex linked genes only from their mothers; Therefore Pm = Pf of the previous generation. Females get their sex linked genes equally from both parents; Therefore Pf = the mean of Pm and Pf in the previous generation. Using primes to indicate the progeny generation.

P1m = Pf
Pf = ½ (Pm + Pf)
Difference between the 2 frequencies in the sexes

P1f – P1m = ½ (Pm + Pf) – Pf
= ½ Pm + ½ Pf – Pf – Pf = ½ Pm – ½ Pf
= -½ (Pf – Pm) = half the difference in previous generation but in the other direction.

Therefore distribution of genes between 2 sexes osatlates but the difference is halved in successive generations as the population rapidly approaches eqm in which the frequencies of the 2 sexes are equal.