# Notes on Hardy-Weinberg

## Example data

| Genotype | $MM$ | $MN$ | $MN$ |
|---|---|---|---|
| No. observed | 363 | 634 | 282 |

## Frequency calculations

$$\begin{aligned}
\text{Genotype frequency} &= \frac{\text{Count of genotypes in sample}}{\text{Number of individuals in sample}} \\
\text{Freq}(MM) &= \frac{363}{363 + 634 + 282} \\
&= 0.28
\end{aligned}$$

$$\begin{aligned}
\text{Allele frequecy} &= \frac{\text{Count of alleles in sample}}{\text{Number of alleles in sample}} \\
\text{Freq}(M) &= \frac{2(363) + 634}{2(363) + 2(634) + 2(282)} \\
&= 0.53
\end{aligned}$$

### Symbolic calculations

| Genotype | $MM$ | $MN$ | $NN$ |
|---|---|---|---|
| Frequency | $u$ | $v$ | $w$ |
| No. observed | $uN$ | $vN$ | $wN$ |
| Sample size | | N | |

$u$, $v$, and $w$ are symbols for the genotype frequencies of $MM$, $MN$, and $NN$, respectively. Let's use $p$ to symbolize Freq($M$). Then

$$\begin{aligned} p &= \frac{2(uN) + vN}{2(uN) + 2(vN) + 2(wN)} \\ &= \frac{2u + v}{2(u + v + w)} \\ &= u + \frac{v}{2} \end{aligned}$$

# Mating table

| Mating | Frequency | Offspring genotype $MM$ | $MN$ | $NN$ |
|---|---|---|---|---|
| $MM \times MM$ | $u^2$ | 1 | | 0 |
| $MM \times MN$ | $uv$ | $\frac{1}{2}$ | $\frac{1}{2}$ | 0 |
| $MM \times NN$ | $uw$ | 0 | 1 | 0 |
| $MN \times MM$ | $vu$ | $\frac{1}{2}$ | $\frac{1}{2}$ | 0 |
| $MN \times MN$ | $v^2$ | $\frac{1}{4}$ | $\frac{1}{2}$ | $\frac{1}{4}$ |
| $MN \times NN$ | $vw$ | 0 | $\frac{1}{2}$ | $\frac{1}{2}$ |
| $NN \times MM$ | $wu$ | 0 | 1 | 0 |
| $NN \times MN$ | $wv$ | 0 | $\frac{1}{2}$ | $\frac{1}{2}$ |
| $NN \times NN$ | $w^2$ | 0 | 0 | 1 |

Let's use $u'$ to symbolize the frequency of the $MM$ genotype in the offspring. If we assume that the four conditions necessary for Hardy-Weinberg apply, we can calculate $u'$ as follows:

$$\begin{aligned} u' &= u^2(1) + uv\tfrac{1}{2} + vu\left(\tfrac{1}{2}\right) + v^2\left(\tfrac{1}{4}\right) \\ &= u^2 + uv + \frac{v^2}{4} \\ &= \left(u + \frac{v}{2}\right)^2 \\ &= p^2 \end{aligned}$$

# An application

Theodosius Dobzhansky studied variation in *Drosophila pseudoobscura*. In one sample from southern California, he found the following genotypes:

| Genotype | $ST/ST$ | $ST/CH$ | $CH/CH$ |
|---|---|---|---|
| No. observed | 57 | 164 | 29 |

$$\begin{aligned} \text{Freq}(ST) = p &= \frac{2(57) + 164}{2(57) + 2(164) + 2(29)} \\ &= 0.44 \\ \text{Freq}(CH) = q &= 1 - 0.44 = 0.56 \end{aligned}$$

If genotypes were in Hardy-Weinberg proportions we would expect to see

| Genotype | $ST/ST$ | $ST/CH$ | $CH/CH$ |
|---|---|---|---|
| No. expected | $(0.44^2)(250)$ | $2(0.44)(0.56)(250)$ | $(0.56^2)(250)$ |
| | 78 | 124 | 48 |

There are *more* heterozygotes observed than expected, suggesting that heterozygotes may be more likely to survive than homozygotes.