



Herausgegeben von der Österreichischen Mathematischen Gesellschaft
<http://www.oemg.ac.at/Mathe-Brief> — mathe-brief@oemg.ac.at

HARDY-WEINBERG-GESETZ

Die Populationsgenetik befasst sich mit der genetischen Zusammensetzung von Populationen. Diese wird von zahlreichen genetischen, ökologischen und evolutionären Kräften gesteuert, wie etwa Mutation, Selektion oder Migration.

Ohne jeglichen Einfluss dieser Kräfte, besagt das sogenannte Hardy-Weinberg-Gesetz, dass die genetische Beschaffenheit einer Population von Generation zu Generation stets erhalten bleibt. Das Gesetz wurde im Jahr 1908 von dem britischen Mathematiker *Godfrey H. Hardy* (1877–1947) und dem deutschen Mediziner *Wilhelm Weinberg* (1862–1937) unabhängig voneinander entwickelt.

Grundlagen der Genetik. Seit den 1940er Jahren ist bekannt, dass die DNA die Trägerin der Erbinformation darstellt. Bei Menschen, Tieren oder Pflanzen, ist die DNA in Form von *Chromosomen* im Zellkern gespeichert. Dabei variiert die Anzahl der Chromosomen von Spezies zu Spezies. In den meisten Zellen liegen sie im Zellkern als Chromosomenpaare vor, wobei sich die Chromosomen eines Paares in ihrer Grundstruktur ähneln. Im Falle sexueller Fortpflanzung stammt jeweils eines der Chromosomen von der Mutter und eines vom Vater. Da der Chromosomensatz im Allgemeinen doppelt vorkommt, bezeichnet man diese Zellen auch als *diploid*.

Für die Ausbildung von Merkmalen, wie der Augenfarbe oder der Blutgruppe, sind Gene verantwortlich. Ein Gen ist ein zusammenhängender Bereich der DNA. Dabei kann ein Gen mehrere Zustandsformen, sogenannte *Allele*, besitzen. Im einfachsten Fall hängt die Ausprägung eines Merkmals nur von zwei Allelen ab, die sich an entsprechenden Positionen eines Chromosomenpaares befinden.

Bezeichnen wir mit $\mathcal{A}_1, \dots, \mathcal{A}_n$ die paarweise verschiedenen, vorkommenden Allele in der Population, so beschreibt der *Genotyp* eines Individuums die Zusammensetzung dieser Allele und wird durch das Paar $\mathcal{A}_i\mathcal{A}_j$ bestimmt. Das äußere Erscheinungsbild des Merkmals wird dann als *Phänotyp* bezeichnet. Der Genotyp kann entweder *homozygot* $\mathcal{A}_i\mathcal{A}_i$ oder *heterozygot* $\mathcal{A}_i\mathcal{A}_j$, $i \neq j$, sein. Die Genotypen $\mathcal{A}_i\mathcal{A}_j$ und $\mathcal{A}_j\mathcal{A}_i$ lassen sich im Allgemeinen nicht unterscheiden, d.h., welches der Allele vom Vater und welches von der Mutter stammt ist nebensächlich. Trotzdem verwenden wir im Folgenden die Konvention, dass bei $\mathcal{A}_i\mathcal{A}_j$ das Allel \mathcal{A}_i von der Mutter und das Allel \mathcal{A}_j vom Vater stammt.

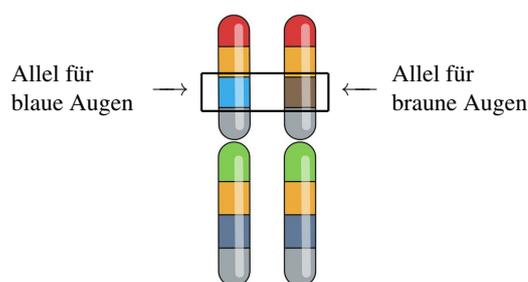


ABBILDUNG 1. Gen für die Augenfarbe auf einem Chromosomenpaar, vgl. [4]

Hardy-Weinberg-Gesetz. Wir betrachten eine sehr große diploide Population mit diskreten, nicht überlappenden Generationen, wie es beispielsweise bei vielen Pflanzen oder Insekten der Fall ist. Ausgeschlossen ist hierbei der Einfluss von Faktoren wie Mutation, Selektion oder Migration. Die Vermehrung erfolgt durch eine zufällige Paarung, d.h. die Partnerwahl ist beliebig.

Ausgehend von einem Gen auf einem Chromosomenpaar, an dem die paarweise verschiedenen Allele $\mathcal{A}_1, \dots, \mathcal{A}_n$ auftreten können, bezeichnen wir mit p_i die relative Häufigkeit des Allels \mathcal{A}_i und mit P_{ij} die relative Häufigkeit des geordneten Genotyps $\mathcal{A}_i\mathcal{A}_j$ in der Population. Da wir zwischen den Genotypen $\mathcal{A}_i\mathcal{A}_j$ und $\mathcal{A}_j\mathcal{A}_i$ nicht unterscheiden, gilt

$$(1) \quad P_{ij} = P_{ji}.$$

Die relative Häufigkeit des ungeordneten heterozygoten Genotyps $\mathcal{A}_i\mathcal{A}_j$ ist durch $P_{ij} + P_{ji} = 2P_{ij}$ gegeben. Da wir von einer sehr großen Population ausgehen, setzen wir relative Häufigkeit mit Wahrscheinlichkeit gleich.

Satz (Hardy-Weinberg-Gesetz).

- (i) Die Allelhäufigkeiten p_i , $i \in \{1, \dots, n\}$, bleiben von Generation zu Generation konstant.
- (ii) Die Genotyphäufigkeit des Genotyps $\mathcal{A}_i\mathcal{A}_j$, $i, j \in \{1, \dots, n\}$, ist ab der nächsten Generation durch $p_i p_j$ gegeben.

Beweis. Zunächst bestimmen wir eine Formel zur Berechnung der Allelfrequenzen p_i für alle $i \in \{1, \dots, n\}$. Ein beliebiges Allel \mathcal{A}_i stammt mit Wahrscheinlichkeit $1/2$ von der Mutter und mit Wahrscheinlichkeit $1/2$ vom Vater ab. Im ersten Fall hat der Genotyp die Form $\mathcal{A}_i\mathcal{A}_j$ und im zweiten Fall $\mathcal{A}_j\mathcal{A}_i$ mit $j \in \{1, \dots, n\}$. Damit ergibt sich für die Allelhäufigkeit p_i des Allels \mathcal{A}_i

$$(2) \quad p_i = \frac{1}{2} \left(2P_{ii} + \sum_{j \neq i} P_{ij} + \sum_{j \neq i} P_{ji} \right) \stackrel{(1)}{=} P_{ii} + \frac{1}{2} \cdot 2 \cdot \sum_{j \neq i} P_{ij} = \sum_j P_{ij}.$$

Um Aussage (ii) zu beweisen, bestimmen wir zunächst die Genotyphäufigkeit des ungeordneten Genotyps $\mathcal{A}_i\mathcal{A}_j$ in der nachfolgenden Generation und bezeichnen diese mit P'_{ij} . Dafür ermitteln wir, aus welchen Paarungen der Genotyp $\mathcal{A}_i\mathcal{A}_j$ entstammen kann. Wir unterscheiden im Folgenden zwischen dem homozygoten und dem heterozygoten Genotyp.

Die möglichen Paarungen aus denen der homozygote Genotyp $\mathcal{A}_i\mathcal{A}_i$ hervorgehen kann, lassen sich durch Vernachlässigung des Geschlechts und der Ordnung der Genotypen in vier Kategorien unterteilen. In Tabelle 1 sind die jeweiligen Paarkombinationen und deren Paarungswahrscheinlichkeiten zusammengefasst, wobei $k, l \in \{1, \dots, n\}$, $k \neq l$ und $k, l \neq i$. Zusätzlich lässt sich die bedingte Wahrscheinlichkeit, dass nach der Paarung Nachkommen mit Genotyp $\mathcal{A}_i\mathcal{A}_i$ existieren, gemäß den Mendelschen Regeln ablesen: siehe Tabelle 1.

$a \times b$	$P(a \times b)$	$P(\mathcal{A}_i\mathcal{A}_i \mid a \times b)$
$\mathcal{A}_i\mathcal{A}_i \times \mathcal{A}_i\mathcal{A}_i$	P_{ii}^2	1
$\mathcal{A}_i\mathcal{A}_k \times \mathcal{A}_i\mathcal{A}_i$	$4P_{ik}P_{ii}$	1/2
$\mathcal{A}_i\mathcal{A}_k \times \mathcal{A}_i\mathcal{A}_k$	$4P_{ik}^2$	1/4
$\mathcal{A}_i\mathcal{A}_k \times \mathcal{A}_i\mathcal{A}_l$	$8P_{ik}P_{il}$	1/4

TABELLE 1. Paarungen für den homozygoten Genotyp

Als Beispiel für die Berechnung der Paarungswahrscheinlichkeiten $P(a \times b)$ wird hier die Kombination $\mathcal{A}_i\mathcal{A}_k \times \mathcal{A}_i\mathcal{A}_i$ gewählt. Die Wahrscheinlichkeit, dass sich ein Individuum mit ungeordnetem Genotyp $\mathcal{A}_i\mathcal{A}_k$ mit einem Individuum mit Genotyp $\mathcal{A}_i\mathcal{A}_i$ vermehrt, ist $4P_{ik}P_{ii}$. Denn wegen $P(\mathcal{A}_i\mathcal{A}_k) = 2P_{ik}$, $P(\mathcal{A}_i\mathcal{A}_i) = P_{ii}$ und weil $\mathcal{A}_i\mathcal{A}_k$ ($\mathcal{A}_i\mathcal{A}_i$) sowohl männlich als auch weiblich sein kann, folgt in Summe $P(\mathcal{A}_i\mathcal{A}_k \times \mathcal{A}_i\mathcal{A}_i) = 2 \cdot 2P_{ik} \cdot P_{ii} = 4P_{ik}P_{ii}$. In ähnlicher Weise lassen sich auch die anderen Wahrscheinlichkeiten berechnen.

Damit können wir die Genotyphäufigkeit P'_{ii} nach dem Gesetz der totalen Wahrscheinlichkeit berechnen

$$\begin{aligned}
 P'_{ii} &= \sum_{a \times b} P(\mathcal{A}_i\mathcal{A}_i | a \times b)P(a \times b) = P_{ii}^2 + 2P_{ii} \sum_{k \neq i} P_{ik} + \sum_{k \neq i} P_{ik}^2 + 2 \sum_{k \neq i} P_{ik} \sum_{\substack{l < k \\ l \neq i}} P_{il} \\
 (3) \quad &= P_{ii}^2 + 2P_{ii} \sum_{k \neq i} P_{ik} + \sum_{k \neq i} P_{ik} (P_{ik} + \sum_{\substack{l \neq k \\ l \neq i}} P_{il}) = (P_{ii} + \sum_{k \neq i} P_{ik})(P_{ii} + \sum_{l \neq i} P_{il}) = (\sum_k P_{ik})^2 \stackrel{(2)}{=} p_i^2.
 \end{aligned}$$

Die möglichen Paarungen, aus denen der ungeordnete heterozygote Genotyp $\mathcal{A}_i\mathcal{A}_j$ hervorgehen kann, lassen sich in fünf Kategorien unterteilen. In Tabelle 2 sind diese aufgelistet, wobei $k, l \in \{1, \dots, n\}$ und $k \neq i, l \neq j, (k, l) \neq (j, i)$. Die einzelnen Wahrscheinlichkeiten lassen sich ähnlich wie im homozygoten Fall berechnen (siehe Tabelle 2).

Analog folgt nach dem Gesetz der totalen Wahrscheinlichkeit für den ungeordneten heterozygoten Genotyp $\mathcal{A}_i\mathcal{A}_j$

$$\begin{aligned}
 2P'_{ij} &= \sum_{a \times b} P(\mathcal{A}_i\mathcal{A}_j | a \times b)P(a \times b) = 2P_{ii}P_{jj} + 2P_{jj} \sum_{k \neq i} P_{ik} + 2P_{ii} \sum_{l \neq j} P_{jl} + 2P_{ij}^2 + 2 \sum_{k \neq i} P_{ik} \sum_{\substack{l \neq j \\ (k,l) \neq (j,i)}} P_{jl} \\
 &= 2P_{ii}P_{jj} + 2P_{jj} \sum_{k \neq i} P_{ik} + 2P_{ii} \sum_{l \neq j} P_{jl} + 2 \sum_{k \neq i} P_{ik} \sum_{l \neq j} P_{jl} \\
 (4) \quad &= 2P_{ii}(P_{jj} + \sum_{l \neq j} P_{jl}) + 2 \sum_{k \neq i} P_{ik} (P_{jj} + \sum_{l \neq j} P_{jl}) = 2(P_{ii} + \sum_{k \neq i} P_{ik})(P_{jj} + \sum_{l \neq j} P_{jl}) \stackrel{(2)}{=} 2p_i p_j.
 \end{aligned}$$

Folglich erhalten wir aus (3) – (4) für alle $i, j \in \{1, \dots, n\}$

$$(5) \quad P'_{ij} = p_i p_j.$$

$a \times b$	$P(a \times b)$	$P(\mathcal{A}_i\mathcal{A}_j a \times b)$
$\mathcal{A}_i\mathcal{A}_i \times \mathcal{A}_j\mathcal{A}_j$	$2P_{ii}P_{jj}$	1
$\mathcal{A}_i\mathcal{A}_k \times \mathcal{A}_j\mathcal{A}_j$	$4P_{ik}P_{jj}$	1/2
$\mathcal{A}_i\mathcal{A}_i \times \mathcal{A}_j\mathcal{A}_l$	$4P_{ii}P_{jl}$	1/2
$\mathcal{A}_i\mathcal{A}_j \times \mathcal{A}_i\mathcal{A}_j$	$4P_{ij}^2$	1/2
$\mathcal{A}_i\mathcal{A}_k \times \mathcal{A}_j\mathcal{A}_l$	$8P_{ik}P_{jl}$	1/4

TABELLE 2. Paarungen für den heterozygoten Genotyp

Analog lässt sich (5) auf alle weiteren Generationen ausweiten, womit Aussage (ii) folgt. Aus (2) und (5) folgt weiters, dass

$$p'_i = \sum_{j=1}^n P'_{ij} = p_i \sum_{j=1}^n p_j = p_i,$$

woraus wir auf Aussage (i) schließen. □

Solche idealen Populationen, wie hier angenommen, treten in der Realität nicht auf, denn es spielen Kräfte wie Mutation, Selektion oder Migration bei der Entwicklung einer Population und deren genetische Variation eine entscheidende Rolle. Trotzdem ist das Hardy-Weinberg-Gesetz die Basis zahlreicher Überlegungen und Ergebnissen in der Populationsgenetik.

Amina Haznadarevic und Gabriela Schranz-Kirlinger (TU Wien)

LITERATUR

- [1] R. Bürger. *The Mathematical Theory of Selection, Recombination, and Mutation*. Wiley, 2000.
- [2] A. Haznadarevic. *Selektion einer diploiden Population an einem autosomalen Genort*. Bachelorarbeit TU Wien, 2020.
- [3] J. Hofbauer and K. Sigmund. *Evolutionary Games and Population Dynamics*. Cambridge Univ. Press, 1998.
- [4] Nagwa. Lesson explainer: Inherited disorders. <https://www.nagwa.com/en/explainers/153127372129/>. letzter Zugriff: 4.6.2024.
- [5] J. Prüss, R. Schnaubelt, and R. Zacher. *Mathematische Modelle in der Biologie : Deterministische homogene Systeme*. Mathematik kompakt. Birkhäuser, 2008.