

Ihre Namen: _____

Übung 1: Populationsgenetik 1, HWE und genetische Drift

Teil 1: Mendelsche Vererbung

1) Schwarze Fellfarbe bei Pferden ist weitgehend bestimmt durch ein rezessives Allel des A-Locus. AA- bzw. Aa-Pferde haben nichtschwarzes Fell, z.B. braun, während aa-Pferde schwarzes Fell aufweisen. Vor ein paar Jahren fragte ein Leser der Internet-Newsgroup rec.equestrian, warum es relativ wenige schwarze Pferde in der Araber-Rassengruppe gibt. Eine Antwort war: „Schwarz ist eine seltene Farbe, weil das schwarze Allel rezessiv ist. Mehr Araber sind braun oder grau, weil diese Farben dominant sind.“ Was ist falsch an dieser Erklärung?

Teil 2: Hardy-Weinberg-Gleichgewicht

Dieser Teil dient dem Verständnis des Hardy-Weinberg-Gleichgewichts und dessen Anwendung.

2) In Menschen kodiert der COL1A1-Locus für ein bestimmtes, in Knochen vorkommendes Kollagen. Das normale Allel an diesem Locus wird mit S bezeichnet. Ein anderes Allel s wird in Verbindung gebracht mit reduzierter Knochenmineraldichte und erhöhtem Risiko von Frakturen bei Frauen nach der Menopause sowohl in ss- als auch Ss-Individuen. Einer Studie aus den Niederlanden zufolge zeigten von 1333 Individuen 895 den SS-Genotyp, 395 den Ss-Genotyp und 43 waren vom Genotyp ss.

a) Entspricht dieser Locus dem Hardy-Weinberg-Gleichgewicht? Führen Sie hierzu den in der Vorlesung besprochenen Chi-Quadrat-Test durch. Als Chi-Quadrat Wert benutzen wir die Irrtumswahrscheinlichkeit 0,05 und 1 Freiheitsgrad, d.h. die Nullhypothese wird abgelehnt wenn der Chi-Quadrat-Wert größer ist als 3,84.

b) Was sagt das Ergebnis über die Auswirkung des s Allels auf die Fitness der Trägerinnen aus?

c) Welche Information benötigen Sie, um sagen zu können, ob die Allele in der nächsten Generation im Hardy-Weinberg-Gleichgewicht sein werden?

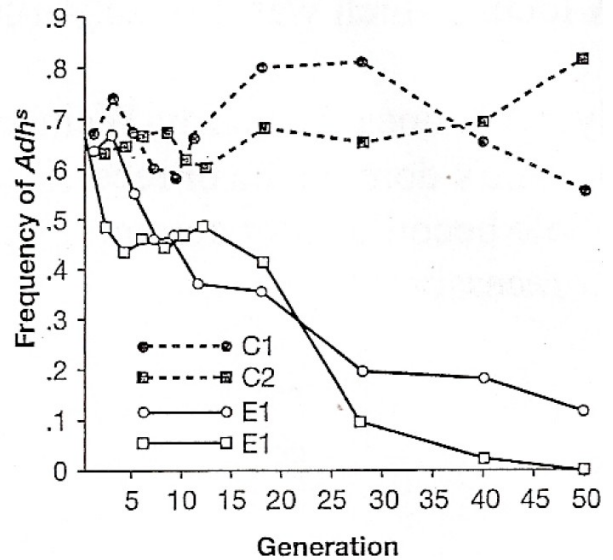
3) Das Hardy-Weinberg-Prinzip lässt sich auf Loci mit beliebig vielen Allelen erweitern. Die erwarteten Häufigkeiten der Genotypen berechnen sich weiterhin als p^2 für Homozygote und $2pq$ für Heterozygote. Betrachten wir hierzu eine Heuschreckenpopulation. In einer Stichprobe finden Sie am A -Locus vier verschiedene Allele. Die Häufigkeiten der Allele A_1 , A_2 , A_3 und A_4 sind $p_1 = 0,6$, $p_2 = 0,2$, $p_3 = 0,15$ bzw. $p_4 = 0,05$.

a) Berechnen Sie unter Annahme eines Hardy-Weinberg-Gleichgewichts den erwarteten Anteil jedes der 10 möglichen Genotypen. (Z.B. sollte der Anteil von A_2A_3 0,1 betragen.)

3b) Zeigen Sie, dass alle Heterozygoten insgesamt 57,5% der Population ausmachen.

3c) Wie viele Exemplare, die heterozygot für das Allel A_4 sind, würden Sie in einer Stichprobe von 100 Heuschrecken erwarten? Wie viele sollten homozygot A_4A_4 sein?

4) In der Abbildung rechts sehen Sie ein Beispiel dafür, wie die Häufigkeit eines bestimmten Allels (Adh^S) in Fruchtfliegen bei Zugabe von Ethanol zum Futter stetig abnimmt (E1). Dies liegt daran, dass Fliegen mit dem Adh^S -Allel besonders schlecht Alkohol abbauen können und somit einen Fitness-Nachteil haben. In Kontroll-Populationen ohne Ethanol würden wir solche Änderungen nicht erwarten. Sehen Sie sich die unselektierten Kontrollen (C1, C2) genauer an: Die Häufigkeit des Adh^S -Allels verändert sich geringfügig, steigt und fällt wieder mit der Zeit. Welche Annahme des Hardy-Weinberg-Modells ist hier höchstwahrscheinlich verletzt worden? Würde dieses Experiment wiederholt, welche Änderung in der Versuchsplanung würde diese Abweichung vom Hardy-Weinberg-Gleichgewicht reduzieren?



Teil 2: Genetische Drift

Im zweiten Teil der heutigen Übungen wollen wir eine Voraussetzung für das Hardy-Weinberg-Gleichgewicht verletzen: Die unendliche Populationsgröße. Sobald Populationen eine endliche Größe haben, findet genetische Drift statt. Um den Effekt von genetischer Drift auf Variabilität innerhalb einer Population besser zu verstehen, benutzen wir das aus der Vorlesung bekannte Wright-Fisher-Modell.

Manuelle Simulation eines Wright-Fisher-Modells mit „Würfeln“ in R

In dieser Übung wollen wir ein Wright-Fisher-Modell "auswürfeln". Die Zahl der Seiten des Würfels entspricht dabei der Anzahl an Individuen in einer Population. Damit lässt sich die Zufallspaarung, wie sie in diesem Modell von einer Generation zur nächsten angenommen wird, simulieren.

Anstatt tatsächliche Würfel zu benutzen, werden wir Zufallszahlen in R generieren. Dies beschleunigt den Vorgang erheblich. Um in R zu würfeln, können Sie folgenden Befehl benutzen: `sample(1:6,6,replace=1)`.

Dies zieht zufällig Zahlen von 1 bis 6, und zwar insgesamt 6-mal. `replace=1` bedeutet, dass eine Zahl mehrmals vorkommen darf. Das Ergebnis entspricht 6 Würfeln eines 6-seitigen Würfels.

Indiv.:

Gen20
Gen19
Gen18
Gen17
Gen16
Gen15
Gen14
Gen13
Gen12
Gen11
Gen10
Gen9
Gen8
Gen7
Gen6
Gen5
Gen4
Gen3
Gen2
Gen1

	1	2	3	4	5	6
Gen20						
Gen19						
Gen18						
Gen17						
Gen16						
Gen15						
Gen14						
Gen13						
Gen12						
Gen11						
Gen10						
Gen9						
Gen8						
Gen7						
Gen6						
Gen5						
Gen4						
Gen3						
Gen2						
Gen1						

Simulieren wir hiermit nun eine Wright-Fisher-Population. Betrachten Sie die Tabelle oben. Jede Zeile stellt eine Generation Ihrer Population dar. Die Spaltenzahl entspricht der Anzahl an Individuen in Ihrer Population, sowie der Anzahl an Seiten des Würfels den Sie benutzen.

5a) Beginnen Sie in der untersten Zeile (Generation 1). Sorgen Sie zunächst für „genetische Variabilität“ indem sie zufällig die Hälfte aller Individuen ausschraffieren. Sie erzeugen dadurch zwei Allele (schraffiert und weiß). Beide Allele starten also mit einer Frequenz von 50%. Würfeln Sie nun die Nachkommen in Generation 2 nach dem Wright-Fisher-Prinzip aus: Für jedes Individuum in Generation 2 wird ein Eltern-Individuum aus Generation 1 ermittelt. Beispiel: Sie starten mit Individuum 1 in Generation 2 und würfeln eine 3. Dies bedeutet, dass das Individuum 3 aus Generation 1 dessen Vorfahre ist. Der Vorfahre vererbt sein genetisches Material (also sein Allel) an seinen Nachkommen, d.h. falls es schraffiert ist, ist das Individuum 1 in Generation 2 ebenfalls schraffiert. Wiederholen Sie diesen Vorgang für alle Individuen der Generation 2. Dabei kann es passieren, dass einige Individuen aus Generation 1 keine Nachkommen produzieren, andere aber dafür sogar zwei oder mehr. Sobald Sie Generation 2 komplett ausgewürfelt haben, wiederholen Sie den Vorgang für Generation 3. D.h. hier würfeln Sie für alle Individuen der Generation 3 die jeweiligen Vorfahren aus Generation 2 aus. Auch hier wird wieder das Allel von Vorfahr an Nachfahr vererbt. Wiederholen Sie diesen Vorgang für mehrere Generationen und füllen Sie die Tabelle aus bis eines der Allele fixiert ist oder Sie Generation 20 erreicht haben.

5b) Können Sie nach 20 Generationen immer noch beide Allele in der Population beobachten? Wie haben sich die Frequenzen verändert? Vergleichen Sie Ihr Ergebnis mit anderen Gruppen.

In einem Wright-Fisher-Modell hängt die Veränderung der Frequenz eines Allels von einer Generation zur nächsten von der Frequenz in der Ausgangsgeneration und der Größe der Population ab. Eine mathematische Analyse zeigt nämlich, dass diese Veränderung *binomialverteilt* ist. In jeder Generation hat ein Allel die Möglichkeit seine Frequenz beizubehalten oder sie zu ändern. Im obigen Beispiel starten wir mit drei schraffierten Allelen und können in der zweiten Generation, je nachdem wie die Würfel fallen, null bis sechs schraffierte Allele erhalten. Die jeweiligen Wahrscheinlichkeiten für diese sieben Fälle folgen einer Binomialverteilung, wobei die Beibehaltung von genau drei Allelen der wahrscheinlichste Ausgang ist.

Wir können diese Erkenntnis nutzen um nun auch Populationen mit sehr vielen Individuen zu simulieren. Wir müssen nicht mehr jedes Individuum einzeln "auswürfeln", sondern können in jeder Generation direkt Zufallszahlen aus einer Binomialverteilung ziehen. In R geschieht dies mit der Funktion `rbinom`.

5c) Wir wollen den Frequenzverlauf eines Allels in Populationen mit 10, 100 und 1000 Individuen verfolgen. Benutzen Sie hierzu die untere Tabelle auf der nächsten Seite.

Wir führen für alle drei Populationsgrößen jeweils **zwei** Simulationen durch. Sie beginnen in Generation 1 jeweils mit der Frequenz eines Allels von 50%. Um nun die Frequenz in der nächsten Generation z.B. für eine Population der Größe 10 zu erhalten führen Sie die Funktion `rbinom(1,10,0.5)` aus. Der erste Parameter besagt, dass Sie ein einziges Mal eine Zufallszahl ziehen wollen, der zweite steht für die Populationsgröße und der dritte für die Frequenz des Allels. Als Ergebnis erhalten Sie die Anzahl and Individuen die das untersuchte Allel in der nächsten Generation tragen. Erhalten sie z.B. eine 4, so hat sich die Frequenz von 0.5 auf 0.4 (4 von 10 Individuen) erniedrigt. Um die Frequenz in Generation 3 zu erhalten führen Sie `rbinom` erneut aus, jedoch diesmal mit der neuen Ausgangsfrequenz: `rbinom(1,10,0.4)`. So können Sie die Veränderung der Frequenz nach und nach simulieren, wobei in jeder Generation die veränderte Frequenz im dritten Parameter von `rbinom` angepasst werden muss. Füllen Sie so die unten

8d) Wiederholen Sie die Simulationen, setzen Sie aber nun $N = 50$. Zählen Sie wieder, wie oft das A-Allel fixiert und sammeln Sie Ihre Ergebnisse in der Tabelle rechts.

Durchgang	Startfrequenz von A		
	$p = 0,2$	$p = 0,5$	$p = 0,7$
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
Mittelwert			

8e) Tragen Sie Ihre Mittelwerte in die Excel-Tabelle des Dozenten ein und notieren Sie die Mittelwerte über alle Gruppe hier:

$p=0,2$:

$p=0,5$:

$p=0,7$:

8f) Haben sich die Fixierungswahrscheinlichkeiten von A geändert oder sind sie mehr oder weniger gleichgeblieben? Was für einen Effekt scheint die Populationsgröße auf die Fixierungswahrscheinlichkeit des hier simulierten A Allels zu haben?

Bis jetzt haben wir mit Allelfrequenzen gearbeitet, die relativ hoch waren. Variabilität innerhalb einer Population entsteht jedoch indem eine neue Mutation (oder neues Allel) zunächst in einem einzelnen Individuum entsteht und sich dann ausbreitet. Hier betrachten wir wieder nur das Schicksal des A-Allels, nicht das vom a-Allel.

9a) Wie hoch ist die Startfrequenz (p) eines neu entstandenen A Allels in einer Population der Größe $N = 5$ bzw. $N = 50$? (HINWEIS: Beachten Sie, dass wir immer von diploiden Organismen ausgehen!)

$N = 5$:

$N = 50$:

9b) Nach dem, was Sie aus Aufgabe 8 und 9a gelernt haben: In welcher der beiden Populationen hat ein neues A-Allel eher die Chance sich durchzusetzen? Führen Sie jeweils 50 Simulationen (mit 10 Loci) zum Schicksal eines neu entstanden A-Allels für jede der beiden oben genannten Populationen ($N=5$ und $N=50$) durch mit der in Frage 9a bestimmten Anfangsfrequenz p . Notieren Sie in wie viel Prozent der insgesamt 500 Simulationen die neue Mutation jeweils fixiert wird.

$N = 5$:

$N = 50$:

9c) Entspricht das Ergebnis Ihren Erwartungen? Was ist der Grund für die eventuellen Unterschiede?

9d) Sie haben gesehen, dass genetische Drift ohne Zuhilfenahme von Selektion zu Evolution im klassischen Sinne führen kann: Eine neue Mutation entsteht und setzt sich dann innerhalb einer Population oder auch einer ganzen Art durch. Solange keine Selektion im Spiel ist, die Mutation also selektionsneutral ist, spricht man von neutraler Evolution. Was würden Sie intuitiv sagen: Kommt solche neutrale Evolution, also die Fixierung von selektionsneutralen neuer Mutationen (Allelen), eher in kleinen oder in großen Populationen vor? Hinweis: Überlegen Sie wie wahrscheinlich es ist, dass sich eine neue entstandene neutrale Mutation in einer Population durchsetzt. Bedenken Sie aber auch, wie viele neue Mutationen in einer Population jede Generation entstehen!